FENCE LIBRARY

FEB 3 1934

Vol. XVII, No. 1 *PSYCHOLOGICAL REVIEW PUBLICATIONS* FEBRUARY, 1934

Journal of Experimental Psychology

EDITED BY

SAMUEL W. FERNBERGER, UNIVERSITY OF PENNSYLVANIA

†HOWARD C. WARREN, PRINCETON UNIVERSITY (*Review*)
WALTER S. HUNTER, CLARK UNIVERSITY (*Index*)
HERBERT S. LANGFELD, PRINCETON UNIVERSITY (*Monographs*)
EDWARD S. ROBINSON, YALE UNIVERSITY (*Bulletin*)

CONTENTS

Learning without Awareness of What is Being Learned or Intent to Learn It: E. L. THORNDIKE AND R. T. ROCK, JR. 1
Changes in Skin Temperature Following Intense Stimulation: H. HELSON AND L. QUANTIUS .. 20
The Effect of Varying External Conditions on Learning, Retention, and Reproduction: A. H. MASLOW 36
Anticipation and Erroneous Responses: F. H. LUMLEY 48
Punishment by Electric Shock as Affecting Performance on a Raised Finger Maze: M. B. JENSEN 65
The Effect of Punishment During Learning upon Retention: L. W. CRAFTS AND R. W. GILBERT 73
Variations of Psychological Measurements Associated with Increased Vitamin B Complex Feeding in Young Children: E. R. BALKEN AND S. MAURER 85
The Influence of Coffee on the Association Constant: G. REIMAN 93
The Statistical Limen Versus the Average as a Measure of Visual Apprehension: M. A. TINKER 105
Judgment Times for the Method of Constant Stimuli: W. R. CARLSON, R. C. DRIVER AND M. G. PRESTON 113
Analysis of Musical Appreciation by Means of the Psychogalvanic Reflex Technique: M. L. PHARES 119
The Effect of Varying Time Intervals between Acts of Learning with a Note on Proactive Inhibition: A. H. MASLOW 141
The Relative Difficulty of Nonsense Syllables: W. C. F. KRUEGER 145
The Sensory Function of the Non-Preferred Hand: J. M. SMITH 154
Selecting and Timing Apparatus: F. H. LUMLEY 160

PUBLISHED BI-MONTHLY
FOR THE AMERICAN PSYCHOLOGICAL ASSOCIATION
BY THE PSYCHOLOGICAL REVIEW COMPANY
PRINCE AND LEMON STS., LANCASTER, PA.
AND PRINCETON, N. J.

Entered as second-class matter, July 1, 1920, at the post-office at Lancaster, Pa.

PUBLICATIONS

OF THE

AMERICAN PSYCHOLOGICAL ASSOCIATION

EDITED BY

†HOWARD C. WARREN, PRINCETON UNIVERSITY (*Review*)
S. W. FERNBERGER, UNIVERSITY OF PENNSYLVANIA (*J. Exper. Psych.*)
WALTER S. HUNTER, CLARK UNIVERSITY (*Index and Abstracts*)
HENRY T. MOORE, SKIDMORE COLLEGE (*J. Abn. and Soc. Psychol.*)
HERBERT S. LANGFELD, PRINCETON UNIVERSITY (*Monographs*)
EDWARD S. ROBINSON, YALE UNIVERSITY (*Bulletin*)

HERBERT S. LANGFELD, Business Editor

PSYCHOLOGICAL REVIEW

containing original contributions only, appears bi-monthly, January, March, May, July, September, and November, the six numbers comprising a volume of about 540 pages.

PSYCHOLOGICAL BULLETIN

containing critical reviews of books and articles, psychological news and notes, university notices, and announcements, appears monthly (10 numbers), the annual volume comprising about 720 pages. Special issues of the BULLETIN consist of general reviews of recent work in some department of psychology.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

containing original contributions of an experimental character, appears bi-monthly, February, April, June, August, October, and December, the six numbers comprising a volume of about 900 pages.

PSYCHOLOGICAL INDEX

is a compendious bibliography of books, monographs, and articles upon psychological and cognate topics that have appeared during the year. The INDEX is issued annually in June, and may be subscribed for in connection with the periodicals above, or purchased separately.

PSYCHOLOGICAL ABSTRACTS

appears monthly, the twelve numbers and an index supplement making a volume of about 700 pages. The journal is devoted to the publication of non-critical abstracts of the world's literature in psychology and closely related subjects.

PSYCHOLOGICAL MONOGRAPHS

consists of longer researches or treatises or collections of laboratory studies which it is important to publish promptly and as units. The price of single numbers varies according to their size. The MONOGRAPHS appear at irregular intervals and are gathered into volumes of about 500 pages.

JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY

appears quarterly, April, July, October, January, the four numbers comprising a volume of 448 pages. The journal contains original contributions in the field of abnormal and social psychology, reviews, notes and news.

ANNUAL SUBSCRIPTION RATES

Review: \$5.50 (Foreign, \$5.75). **Index:** \$4.00 per volume.
Journal: \$7.00 (Foreign, \$7.25). **Monographs:** \$6.00 per volume (Foreign, \$6.30).
Bulletin: \$6.00 (Foreign, \$6.25). **Abstracts:** \$6.00 (Foreign, \$6.25).
Abnormal and Social: \$5.00 (Foreign, \$5.25). Single copies \$1.50.
Current numbers: **Journal,** \$1.25; **Review,** \$1.00; **Abstracts, 75c; Bulletin, 60c.**

COMBINATION RATES

Review and Bulletin: \$10.00 (Foreign, \$10.50).
Review and J. Exp.: \$11.00 (Foreign, \$11.50).
Bulletin and J. Exp.: \$12.00 (Foreign, \$12.50).
Review, Bulletin, and J. Exp.: \$16.00 (Foreign, \$16.75).
Review, Bulletin, J. Exp., and Index: \$19.00 (Foreign, \$19.75).

Subscriptions, orders, and business communications should be sent to the

PSYCHOLOGICAL REVIEW COMPANY

PRINCETON, NEW JERSEY

Journal of
Experimental Psychology

VOL. XVII, No. 1 FEBRUARY, 1934

LEARNING WITHOUT AWARENESS OF WHAT IS BEING LEARNED OR INTENT TO LEARN IT[1]

BY EDWARD L. THORNDIKE AND ROBERT T. ROCK, JR.

Institute of Educational Research
Teachers College, Columbia University

In *The Fundamentals of Learning* (E. L. Thorndike and others, 1932), evidence was presented that a satisfying after-effect could strengthen the connection which it followed and to which it belonged in cases where the learner did not know what the connection was. This matter is so important for learning in the sphere of emotions, attitude, and conduct, and for general theory, that we have devised and carried out two additional experiments, one on the strengthening of an intellectual tendency, the other on the strengthening of a motor tendency. We may call these the Association Experiment and the Four Quadrants Experiment.

In the former, the subject is instructed as follows: "I shall say 'Ready' and then say a word. You will say the first word or words that you think of. I will record the time in fifths of a second. Then I will say another word. You will say the first word or words that you think of. I will record the time as before. We will continue this. At the end of the experiment, you will receive a money bonus depending

[1] The investigation reported here was part of a general program of work on the fundamentals of learning, supported by a grant from the Carnegie Corporation.

upon how quickly you respond on the average. Also I shall say 'Right' if the word that you say is one of ten or more that we have arbitrarily decided to call right as responses to the word in question. I shall say 'Wrong' if the word that you say is one of ten or more that we have arbitrarily decided to call wrong as responses to the word in question. Sometimes I shall say nothing because the word that you say is in neither our list of Rights nor our list of Wrongs for the word in question. At the end of the experiment, you will receive a money bonus depending on a system of credits for Rights and penalties for Wrongs."

320 or 640 words are then read to him, one at a time. In each case, if his response is clearly due to the sequential connections used in speaking and writing (such as yours → truly, achieve → success, across → the street, absurd → statement, up → stairs) he is rewarded by the announcement 'Right.' If his response is clearly due to connections used in getting the word's meaning (such as yours → mine, achieve → get, across → over, absurd → unreasonable, up → down) he is punished by the announcement of 'Wrong.' Thus for *adverb* any of the following would be called right: and, of, in, of time, of degree, meaning, used, is, was, will be, that, which, etc., etc. Any of the following would be called wrong: verb, adjective, noun, plainly, slowly, quickly. For *achieve* any of the following are right: success, much, little, results, good, good success, more, what, what he desires, etc., etc. Any of these would be called wrong: win, accomplish, get, obtain, do. For *across* the following are right: the, the street, the river, the sea, oceans, deserts (person, child, boy, girl—these being due to hearing *across* as *a cross*), etc., etc. Any of these would be called wrong: near, away, over, sideways, about, toward.

If the experimenter cannot decide immediately which tendency is operating, no announcement is made. As was shown in *The Fundamentals of Learning* (Chapter 14) over ninety percent of the first words that come to mind in the free association experiment are due to one or the other of these two tendencies. But the subject, unless he has chanced to read the book in question, does not know about them.[2]

[2] At the time of the experiments the book was still unpublished.

If he thinks about the experiment at all, he probably thinks of it as a test of his sanity, or complexes, or interests, or speed of reaction, or expects the same words to recur. If he did anything voluntarily, he might repeat to himself the stimulus word and his response in cases where he had it right. He has, however, little time to do this.

To prevent the subjects from inferring that opposites were always or nearly always wrong, certain opposites which might well come by the action of the tendency to operate the connections used in speaking and writing, were always called right. Such were far → near, true → false, good → bad, eat → drink, give → take, hand → foot. To make the administration of this rule easy, *any response* to far, true, good, eat, etc. was called right. There were two such words in each 40.[3]

If the subject became aware of the principle according to which the rewards and punishments were allotted or of the two classes of connections, he would, after a little experimentation, change from a moderate percentage of successes to nearly 100 percent. If the subject increases his percentage of responses that manifest the one tendency gradually, it is evidence that that tendency is strengthened without his being aware that there is any such tendency.

The experiment is a good one because the rewarded tendency and the punished tendency are real, are distinct, and are ready to act, though the subject is ignorant of their existence.

The following special rules were also adopted: -hood, -ness, -ly, -ment, and similar suffixes were called wrong. When the response was classifiable as a suffix but where it was a word in itself, which added to the stimulus word gave a new word, it was counted correct. For instance, to the stimulus word *port*, the response *folio* was sometimes secured. This was counted correct.

Certain responses which could conceivably be of the sequential sort used in speaking and writing but which could also be considered of the sort used in getting a word's meaning

[3] In the case of the first fifteen of the subjects a different rule was used to the same end. One opposite out of each successive forty responses was arbitrarily called right.

were counted wrong arbitrarily. The most frequent of these were: hammer → nail, cry → tears, scissors → cut, cross → street, butter → bread.

In a few instances among the first fifteen, some subject adopted one response which he had found successful and used it for a long series of words. After one happening of this sort, the practice was adopted of calling the last response wrong after the subject had used it five times in a row. In the case of sixteen of the subjects, the rule was made more stringent and any words after the first in such a sequence were called wrong. If the subject used it once in a while, it was counted right later on in the series. An example of this type is the use of the word *and* as the response.

To give a more realistic presentation of the responses and of our classification of them, I quote those to 6 beside, 8 begin, 13 balmy, 19 gentle, 20 gift, 244 among, 251 cheap, 252 defer, 253 decent, and 245 bargain.

Fifteen of the subjects had the former five in their first twenty and the latter in their thirteenth twenty. Sixteen had the former five in their thirteenth twenty and the latter in their first twenty.

Responses to Ten Words Taken as a Sample

1 = speech-writing connections
? = doubtful or other connections
2 = meaning connections

By persons for whom these words occurred in the first twenty of the experiment.

Beside

1. her me (3) myself there
?. chair lake sit
2. above across (2) at near next

Begin

1. here now (2) saying some the experiment
?. began
2. end start (7)

Balmy

1. breeze day ocean summer
?. banana him (omitted)
2. calm clear foolish rainy warm (3) wet

Gentle

1. boy girl man master stroke woman
?. snow
2. angry docile easy kind (2) mild slow soft

Gift

1.
?. book Christmas (6) genius low now
2. giver present (4)

Among

1. bushes friends (1) others people (2) them
?. there
2. between (3) in together with (2) within

Bargain

1. counter (2) sale (5)
?. buy goods price (2)
2. cheap contract money sell trade

Cheap

1. goods wares
?. price (2) year (omitted 2)
2. bargain (2) dear (3) high inexpensive (2) little

Defer

1. a little payment (3) present
?. confer later time (omitted)
2. away postpone (2) put off (2) wait (2)

Decent

1. book clothes company looking
?.
2. good (6) immodest indecent (2) right (2) upright

Responses to Ten Words Taken as a Sample

1 = speech-writing connections
? = doubtful or other connections
2 = meaning connections

By persons for whom these words occurred in the last twenty of the experiment.

Beside

1. brook himself river the lake the river us you
?. people table water
2. close near (4) next

Begin

1. life over story the day the lesson the work work (2)
?.
2. end start (6) (omitted)

Balmy
1. air breeze day (2) evening seas weather
?. light south
2. airy clear (2) nice soft sweet

Gentle
1. child horse (2) man mother (2) person woman
?.
2. calm kind (2) kindly soft (2) sweet swift

Gift
1. of talk to to men
?. Christ Christmas (3) good jewels love
2. giver present (5)

Among
1. friends many neighbors the forest the men (2) the people those us
?. hours yours tree
2. amidst in with

Bargain
1. counter (4) house on the table sale (2) sales with
?. bad basement (illegible)
2. cheap trade

Cheap
1. and not good apples furniture goods (2) house literature or costly woman
?.
2. bargain dear expensive (2) little meager

Defer
1. business payment (8) to (2) our payment
?. from me
2. stalling wait

Decent
1. and true behavior company gentlemen girl (2) man (3) person quite
?.
2. clean good indecent respectable

There is learning. The total number of right responses in the first ten was 80, and, in the last ten, 204. In the first forty it was 358, and in the last forty, 720. Little, if any, of this learning can be ascribed to awareness of the distinction. The detailed records presented in Table 1 do not show sudden changes in the percentage correct. The most suspicious cases are of subjects L and C. L probably made the change without such awareness during trials 41 to 120, his score during trials 1 to 140 by twenties being 4, 4, 11, 15, 16, 18, 19. We have

inspected the responses in order to observe whether, by any reasonable interpretation, we can change the scores so as to make this part of L's record show a more sudden shift and be consistent with learning by awareness of the distinction and deliberate choice of the one sort of responses. By the most favorable interpretations, we are able to make the scores by tens 0, 4, 2, 2, 9, 10, 10, 7, 8, 10, 10, 9, 10. The 7 and 8 are not consistent with such awareness. Moreover, the next 10 (141 to 150) contain three that cannot by any stretch of allowance be regarded as due to speech-writing associations.

In the case of C the first two responses (*it* to *adverb* and *wind* to *achieve*) are due to neither sort of connections, but apparently to the idea 'say some word quickly.' The next one *across* to *pond* was scored wrong, but probably had its source in 'across the pond.' It surely belongs rather to the speech-writing tendency than to the get-to-the meaning tendency. *Absurd* to *ridiculous* is a case of a meaning connection. From then on we have a series of clear speech-writing connections to No. 100 except for film → picture, gift → now, lawn → money (probably lawn was misheard as loan), barren → fertility, dative → in history (probably misheard as dates), and firmament → exists. So the case of C seems to be that of a person who, except for the first two or three random responses, had the one tendency at almost full strength from the beginning. He probably was not aware of it and consequently shifted occasionally as in gift → now, barren → fertility, gas → dangerous. Statistically, he did not learn after trial 2 or at least trial 4.

We have inspected the records of individuals A, I, P and R in a similar manner. A gives no evidence of awareness of the distinction. Even to the end he inserts some responses due to meaning associations. The same is true of I. Her last score of 9 is correct, the wrong one being money → cash, but in the 27th ten she gives quiet → peace, short → wide, and swift → long. In the 26th ten she gives cabbage → cauliflower, earth → heaven, hard → soft, and justice → honor, which nobody would give who was aware of the distinction. The same is true of P. By straining interpretation somewhat

we can make her scores in trials 71 to 80, 81 to 90, 91 to 100 and 101–110 6, 6, 10 and 10, instead of 6, 5, 7, and 9, as in Table 1, but the 7 and 7 next following cannot by any interpretation be made into 10's. They include also → too, annual → yearly, several → two, and sugar → flavoring. Later on we find fast → slow, dative → reason, genitive → neuter, whistle → call.

R is the most plausible case for learning by awareness. Her early learning was not so, but progressed slowly with successive scores of 0, 2, 0, 0, 2, 1, 3, 2, 3, 4, 4, 4. In the next ten, seven were announced as right and *ten → eleven* should properly be called so. In the next ten, the only wrong item is heat → warm; and thereafter we can, by straining interpretation occasionally, attribute all responses to the speech-writing tendency except soft → hard, whistle → noise, fell → feather, adverb → adjective, darling → dear, immoral → bad, whiskey → ale, and simple → difficult. Even so it may be doubted whether a person who was aware of the distinction would say 'whistle → noise' instead of 'whistle a tune,' 'whistle loudly,' or the like, 'darling dear' instead of 'darling baby,' 'darling child,' or the like, and 'simple difficult' instead of 'simple question,' 'simple dress,' or the like.

The learning was then, except possibly in R, not due to awareness of the distinction. Our next precaution must be to learn whether it was due in part to other sorts of awareness, such as notions that 'of' is often right, 'and' is often right, 'it' is often right.

We have examined the responses made by B, G, H, J, M, N, O, S, and CC with this matter in view. M said 'me' as a response to her sixth word (beside), and from time to time thereafter, and 7 of her last 40 responses were 'me.' She also used 'of' more and more frequently in later trials. But these possible insights do not account for all of her learning, or for half of it. The improvement of B, G, H, J, N, O, S, and CC seems in no wise due to such notions. P, whose case has been discussed already in another connection, used 'and,' 'of,' 'to,' 'me,' 'up,' and 'them' much more frequently in

the later tasks than in early ones, and this may have been in part deliberate. If we leave out of account all responses by these words her scores (percents correct) for the first, second, third, thirteenth, fourteenth, and fifteenth twenties change from 35%, 5%, 35%, 65%, 85%, and 85% to 28%, 5%, 28%, 56%, 86% and 81%. The gain including 'and,' 'of,' 'to,' etc. was then hardly greater than the gain without them.

It remains to notice a fact about this experiment which deserves explanation, namely the great variability among individuals in the amount of strengthening of the rewarded tendency. Nearly one-fourth of the persons probably did not learn at all. In one case (C) already mentioned this is because the subject really had a perfect record after task 4 (or possibly after task 2). This was also the case with DD. But in E, F, K, Q, and V, especially in E, F, and V where the rewarded sort of connections occurred five or more times in the first twenty, the failure of the reward to strengthen the tendency should be explained.

One probable influence was emphasis upon speed in the instructions.[4] This emphasis was desirable, in order to ensure the frequent operation of the two tendencies one of which was to be rewarded, and to prevent the subjects from thinking *about* their responses, and from thinking of several words and then selecting one for report. But if a subject was led by the instructions to care little whether the announcement was 'Right' or 'Wrong,' and to pay little or no attention to the announcements, he would, of course, strengthen the rewarded tendency little or not at all.

Another influence which we find in all such experiments is a certain over-complication of the procedure by irrelevant emotional or intellectual attitudes of the subject. It is our opinion that if these subjects had done the experiment naively, without trying to appear intelligent, or to make a good impression on the experimenter, or to solve some system which they imagined was back of the choice of 'Rights' and 'Wrongs,' they probably would have learned. If the tendencies that produce a word or words when a word is heard

[4] "Speed is the main thing. The number of rights versus wrongs is a minor thing."

TABLE 1

NUMBER OF RESPONSES ANNOUNCED AS 'RIGHT' IN SUCCESSIVE TENS FOR INDIVIDUALS A, B, C, ETC.

	A	B	C		E	F	G	H	I	J	K	L	M	N	O		P	Q	R	S	T	U	V	W	X	Y	Z	AA	BB	CC	DD	EE
I.	3	1	6		3	3	1	6	4	3	0	1	2	2	1		1	1	0	1	2	2	5	3	6	2	1	1	1	1	10	5
II.	6	5	6		9	4	2	7	4	2	1	4	3	2	4		6	2	3	3	1	3	8	4	9	3	2	2	1	0	7	5
III.	4	1	8		6	4	2	5	7	0	0	3	0	0	1		0	1	0	1	1	2	8	2	7	0	1	2	0	4	7	5
IV.	5	1	10		3	3	3	4	1	1	1	2	1	1	2		1	1	0	1	2	2	5	1	8	2	0	1	0	5	6	2
V.	7	1	9		4	3	1	7	3	2	2	5	1	4	4		4	3	2	3	2	3	4	2	6	2	2	1	0	3	9	1
VI.	5	3	10		7	7	1	3	4	3	1	8	0	8	7		3	4	0	2	2	0	6	1	6	2	2	0	1	5	9	1
VII.	7	2	9		4	5	3	4	7	6	1	9	2	7	9		4	3	1	2	1	3	5	2	5	0	0	1	0	4	7	2
VIII.	7	0	9		9	4	1	0	6	3	2	7	4	8	6		6	1	2	2	5	3	5	4	7	1	1	2	0	3	8	2
IX.	7	0	9		7	7	2	4	8	6	0	9	4	8	5		5	2	3	1	0	1	5	0	8	0	0	2	0	2	10	2
X.	8	1	6		4	2	3	2	7	5	0	8	2	8	4		7	1	4	0	3	0	3	2	5	2	0	0	0	0	9	1
XI.	9	3	8		5	4	0	5	4	4	1	9	4	6	6		9	2	4	2	1	1	2	2	4	0	0	1	0	2	10	1
XII.	6	6	9		1	2	1	5	5	5	2	8	1	1	4		7	1	4	2	2	2	3	4	4	2	1	1	0	3	9	1
XIII.	7	4	10		7	6	3	8	9	7	1	8	5	5	9		7	3	7	3	4	3	6	4	7	2	3	3	1	5	8	4
XIV.	8	1	10		8	6	3	9	7	8	1	9	4	7	8		10	2	9	1	1	2	7	1	8	1	0	1	1	3	8	1
XV.	9	2	9		6	4	4	10	10	10	0	10	7	8	9		10	1	10	1	0	0	5	5	8	1	1	1	0	4	9	6
XVI.	6	4	10		5	3	6	10	10	8	3	8	7	8	7		9	1	9	0	0	2	8	3	8	1	1	0	0	4	7	2
XVII.	8	8	8		9	5	1	9	9	10	0	10	6	6	10		7	2	10	0	2	2	2	2	10	1	0	2	0	1	5	2
XVIII.	6	0	9		3	4	2	5	8	8	1	10	6	2	4		10	1	8	2	1	2	6	4	6	0	3	1	0	6	8	3
XIX.	8	2	10		4	4	1	5	7	10	0	9	4	4	4		8	0	4	0	0	0	4	1	5	1	2	0	0	3	7	0
XX.	6	2	9		3	1	4	4	3	9	2	9	4	2	3		9	3	9	5	4	2	8	7	9	2	3	1	1	2	10	8
XXI.	6	6	10		6	4	3	6	7	7	3	8	4	9	8		10	1	10	4	2	3	6	5	10	3	0	0	0	3	9	5
XXII.	9	4	10		5	6	5	9	6	8	3	9	3	6	7		8	2	9	2	1	3	4	3	10	4	1	3	0	3	9	4
XXIII.	10	4	10		4	4	1	7	9	9	1	10	2	6	8		6	2	9	3	2	2	5	4	9	3	2	0	2	3	9	3
XXIV.	9	1	10		5	3	1	9	7	10	1	7	1	6	4		9	3	10	1	2	2	8	1	9	1	0	3	1	3	9	5
XXV.	6	3	8		7	3	1	6	10	10	0	9	7	7	5		7	2	5	2	5	2	4	2	6	4	1	2	0	3	9	4
XXVI.	9	6	10		7	7	4	8	8	10	3	8	5	7	9		6	1	9	1	3	0	4	3	7	1	1	1	0	4	8	3
XXVII.	7	6	10		6	7	5	6	8	10	3	10	7	9	6		8	4	7	4	1	2	2	1	9	2	2	1	0	6	8	1
XXVIII.	6	5	10		8	2	2	8	3	10	0	9	6	5	7		9	2	9	5	4	2	5	5	9	2	2	1	0	4	9	2
XXIX.	7	6	10		6	4	5	9	10	7	2	9	9	9	5		8	1	8	2	5	2	7	9	9	1	0	1	0	4	8	4
XXX.	8	3	10		7	4	3	7	9	10	1	8	8	5	5		9	5	10	5	7	4	8	10	10	1	0	2	1	2	10	6

had been allowed to operate undisturbed, the speech-writing tendency would, in our opinion, have been strengthened in all persons in whom it operated at all. But we cannot prove this.

Table 1 presents the facts concerning the number of announcements of Right in each successive ten for each person. Table 2 presents the facts concerning the number of responses in each ten of the first 80 and last 80 which seemed probably due to the speech-writing sort of connections when the subject's responses (which were all recorded by an assistant to the experimenter) were examined later. Obviously the experimenter, who made his judgments as soon as the response was spoken, could not be infallible. Moreover, some of the rules set up to make the experiment more productive deliberately call wrong certain responses probably due to speech-writing connections. Table 1 shows primarily the amount of reward received by the subjects, but also shows roughly the amount of strengthening of the rewarded tendency. Table 2 shows the latter more precisely. In subjects A to O in Table 1, and in all subjects in Table 2, no account is taken of the items which were called right regardless of what the response was or because that item was used to prevent the adoption of the idea "no opposite is ever right."

Fifteen of the subjects (A to O) had the tasks in the order 1 to 320. Sixteen (P to EE) had them in the order 241 to 320, 161 to 240, 81 to 160, 1 to 80.

Subject D is omitted from the records since she is the one who formed the habit of saying 'and,' which was dealt with irregularly in her case. She undoubtedly learned, however.

The Four Quadrants Experiment

The rewarded response in this experiment was to hit within a certain distance from a point *and also in a certain quadrant or direction from the point.* That is, a constant error of direction was rewarded, provided it was not too great in amount. The subjects knew nothing of this, thinking they were rewarded for general precision. The essentials of the

TABLE 2

NUMBER OF RESPONSES ATTRIBUTABLE TO THE SPEECH-WRITING SORT OF CONNECTIONS IN THE FIRST 8 TENS AND LAST 8 TENS

	A	B	C	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	BB	CC	DD	EE	All
I.	4	1	6	3	2	1	7	4	3	0	0	2	2	1	2	1	0	1	2	2	7	3	6	2	0	1	1	1	10	5	80
II.	7	4	7	9	4	1	6	5	2	1	4	2	1	3	7	1	2	3	1	2	9	4	9	3	1	1	1	0	9	5	114
III.	7	0	10	6	4	1	9	6	0	0	2	0	0	1	0	1	0	1	1	2	9	1	8	0	1	2	0	4	8	5	89
IV.	3	1	10	4	3	1	6	2	0	1	2	0	2	1	0	0	0	1	2	1	5	1	9	2	0	1	0	6	9	2	75
V.	9	2	10	6	5	1	6	7	1	2	9	0	5	7	4	3	2	3	2	4	4	2	9	2	2	1	0	4	9	1	122
VI.	8	3	10	7	6	0	1	4	5	1	10	0	6	7	5	2	1	3	3	1	7	2	7	2	2	1	1	5	9	1	120
VII.	7	1	10	8	7	0	0	8	5	2	10	4	9	7	5	4	3	3	2	4	7	3	6	0	0	1	0	5	9	3	133
VIII.	10	0	10	8	7	0	2	6	4	0	6	1	7	8	6	2	2	2	4	3	5	4	7	1	1	2	0	2	9	3	122
	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
XXIII.	9	1	10	6	4	1	10	7	10	2	9	5	7	4	7	2	9	4	1	2	6	5	9	3	2	0	2	5	9	5	156
XXVI.	6	5	9	8	5	0	5	9	9	1	9	6	7	7	10	3	10	1	5	2	10	3	10	2	0	3	1	4	9	6	165
XXV.	9	8	10	9	8	5	8	9	10	2	9	6	9	7	9	2	6	1	5	1	4	3	6	3	1	2	0	3	9	4	168
XXVI.	7	4	10	3	5	3	6	5	8	0	8	9	5	7	6	1	9	2	4	0	5	4	8	0	0	1	0	4	9	3	136
XXVII.	4	9	10	8	2	1	10	7	6	1	10	7	9	9	10	3	7	4	2	1	4	2	9	2	2	1	0	6	8	2	156
XXVIII.	8	5	10	8	5	4	9	9	9	0	9	9	6	6	9	1	9	5	5	2	8	5	9	2	2	1	1	3	10	5	174
XXIX.	9	2	10	10	7	7	7	10	9	3	9	9	9	8	8	2	9	2	5	1	9	10	10	1	0	2	0	4	9	5	186
XXX.	10	5	10	10	8	2	8	9	10	1	9	10	9	7	10	6	10	5	7	3	10	10	10	0	2	1	0	5	10	7	204
First 80	55	12	73	51	38	5	37	42	20	7	43	9	32	35	29	14	10	17	17	19	53	20	61	12	7	10	3	27	72	25	
Last 80	62	39	79	62	44	23	63	65	71	9	72	61	61	55	69	20	69	24	34	12	56	42	71	13	9	11	4	34	73	37	

apparatus and procedure were that the subject put his left forefinger on a point on the under surface of a block, and then tried to put a pencil point on the corresponding point of its top surface (which was hidden from his vision by a screen).

A description of the conduct of the experiment together with the instructions to the subjects is given below. Its plan and execution are due to Dr. R. Rock.

TABLE 3

FOUR QUADRANTS EXPERIMENT. NUMBER 'RIGHT' IN EACH SUCCESSIVE 50 TRIALS FOR EACH INDIVIDUAL

Series	Individual	Number of dots in a 2" by 2" square forming one of the quadrants about the point aimed at, in 16 successive sets of 50 trials															
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
2	1	0	12	16	23	21	22	12	23	13	12	7	6	8	6	7	13
	2	0	0	0	4	1	2	1	0	0	0	0	5	0	0	0	0
	3	1	4	16	8	8	21	19	17	21	32	22	22	33	36	38	28
3	4	7	3	27	24	36	49	23	34	12	30	47	49	46	48	44	44
	5	29	24	28	20	41	30	27	34	38	26	31	39	34	36	31	23
	6	20	36	33	37	34	39	38	35	30	27	27	30	41	47	36	28
	7	27	32	31	31	39	28	20	34	26	28	16	25	29	37	33	32
	8	11	11	42	47	48	48	48	45	30	37	48	49	49	49	49	48
4	9	33	29	34	21	37	32	44	32	46	44	41	42	47	45	47	47
	10	0	11	26	31	20	7	24	22	11	22	20	12	12	16	12	9
	11	1	0	1	0	2	2	0	14	36	32	30	32	16	25	7	14
	12	3	4	2	3	10	14	8	8	14	8	5	1	4	2	2	0
5	13	43	41	38	45	40	40	37	40	46	41	41	39	33	30	25	29
		6	0	0	0	0	0	0	0	39	48	50	45	49	48	48	46
		14	7	13	29	24	17	22	15	38	43	49	41	42	34	41	41
	16	28	31	25	24	21	34	30	30	27	19	17	12	8	8	9	10
6	17	30	29	20	16	13	18	16	33	13	25	18	15	22	32	34	37
	18	1	0	4	6	8	11	6	1	4	3	9	6	6	21	20	38
	19	25	43	45	49	48	47	44	48	28	42	49	44	46	48	48	49
	20	3	8	30	39	35	37	34	35	18	36	31	34	19	5	22	11
	1	19	11	12	30	17	28	22	28	32	33	36	29	31	28	28	12
	2	8	12	7	11	15	14	23	30	22	29	17	32	17	14	28	34
	3	0	2	0	4	4	6	0	2	9	10	8	16	12	22	45	39
	4	24	22	28	32	39	37	43	40	37	49	39	38	38	44	44	39
	25	0	2	9	0	0	0	10	18	25	18	10	15	23	26	24	31
	26	0	3	2	6	12	12	14	14	0	0	3	5	6	5	8	2
	27	3	1	10	0	1	3	0	0	3	2	2	0	0	1	2	0
	28	20	14	28	13	16	15	16	12	2	1	8	1	3	0	3	1
	29	0	1	0	2	0	0	0	3	18	26	7	5	13	36	7	29
	30	19	28	21	24	25	32	32	28	7	14	22	39	45	44	39	40
		1	0	1	0	1	0	0	4	4	0	1	0	0	0	0	0
		9	12	18	11	12	17	11	16	19	31	25	36	14	40	16	21
		2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		13	20	28	41	29	36	43	43	26	20	8	8	0	8	5	12
Average		11.8	13.6	17.5	18.6	19.3	20.5	19.6	21.7	20.4	22.3	21.9	22.7	21.9	24.7	23.6	23.9
Percent		24	27	35	37	39	41	39	43	41	45	44	45	44	49	47	48

The learning is very variable among individuals and within the same individuals as is shown by the detailed scores of Table 3. There can be no doubt, however, that the tendency to make a certain systematic error is in general strengthened. The average percent of dots within the chosen quadrant rises from 24 to 48 from the first to the sixteenth fifty. They run in order, 24, 27, 35, 37, 39, 41, 39, 43, 41, 45, 44, 45, 44, 49, 47, and 48. There was no adequate evidence that a single subject of the thirty-four subjects discovered the system. There are sudden rises in the percentage of 'Right' responses, but there are equally sudden drops. At all events we find learning still present if we omit the records of all individuals who might conceivably have learned by inferring the basis for the announcements of 'Right.' After omitting I. He, and A. Gr., the average percentages for the first and sixteenth fifties are 24 and 45.

Apparatus, and Conduct of the Experiment

The location of each of the ten points aimed at is shown in Fig. 1, which represents the top of the block as the experimenter saw it and also the bottom of the block fitted with ten tasks as the subject saw it in a mirror and felt it with his left forefinger. Figure 2 shows the arrangement of the apparatus as a whole.

Six series of experiments were run during October, November and December 1931.

Series 1.—Quadrant one, from the position of the operator, was chosen arbitrarily as the 'correct' quadrant and the limits for a right response were the boundaries of one inch square having the origin at one corner and the x and y axes as sides. The origin was in each case a point directly above the tack on the lower surface. Misleading preliminary trials had made the one inch limits seem reasonable but with the ten subjects run for one hour each it was found that this area was too small to yield a satisfactory number of right responses. The data from this first series are not included in the data summarized for this experiment.

Series 2, 3, 4 and *5.*—The area within which a response

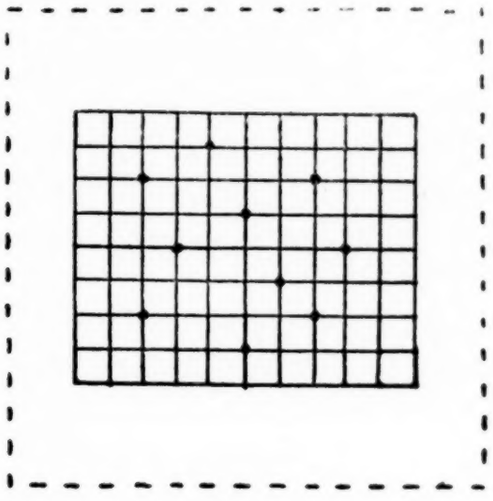


FIG. 1.

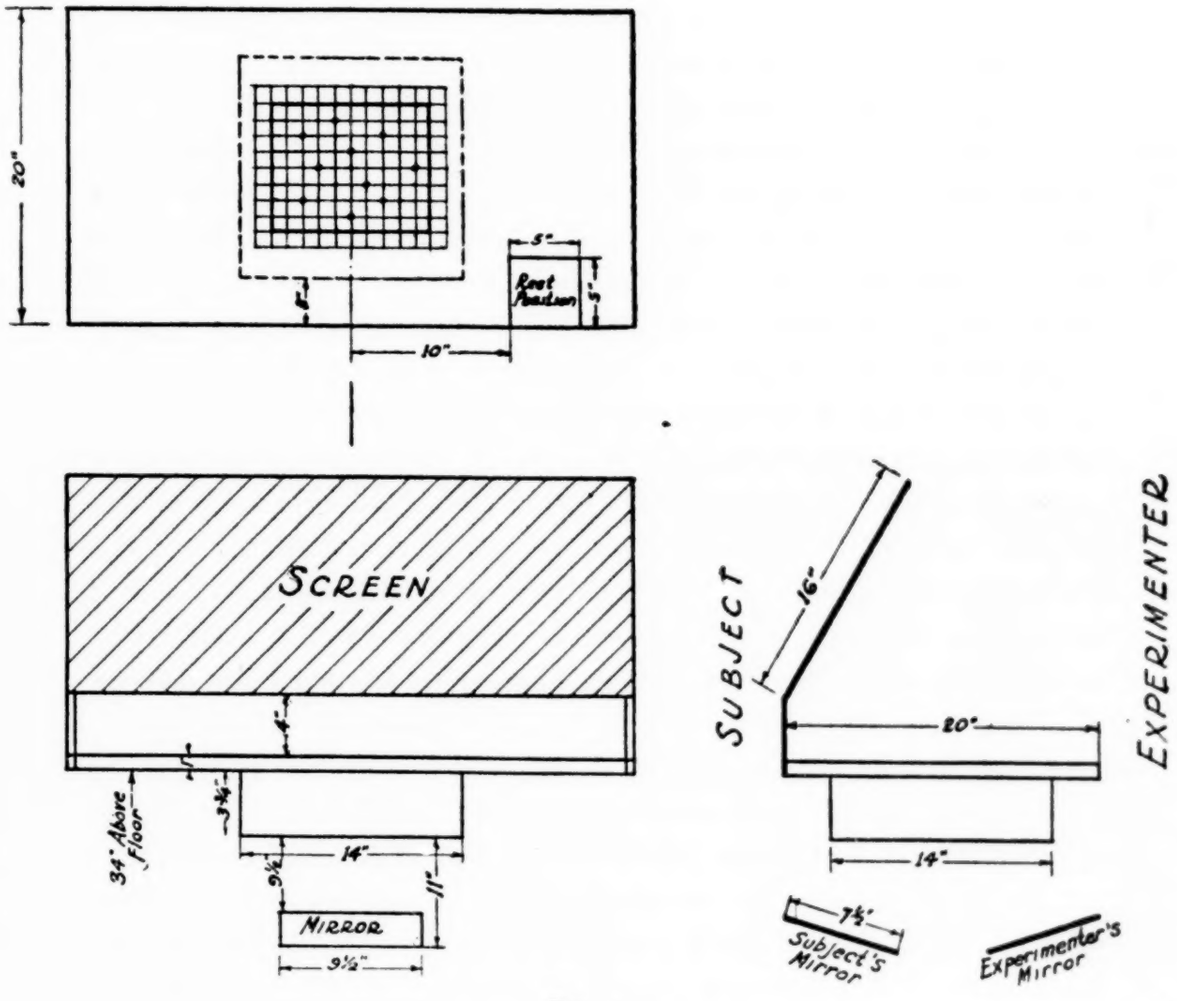

20"
5"
Rest Position
3"
8"
10"
SCREEN
4"
1"
3½"
34" Above Floor
14"
9½"
11"
MIRROR
9½"
SUBJECT
16"
20"
14"
EXPERIMENTER
7½"
Subject's Mirror
Experimenter's Mirror

FIG. 2.

was to be called right was increased for these series to a $2'' \times 2''$ square having the origin at one corner and the x and y axes as two sides. In Series 2 this square was in the third quadrant; in Series 3 in the second quadrant; in Series 4 in the fourth quadrant; and in Series 5 in the first quadrant. In each case the quadrant is named from the operator's position.

Series 6.—In the preceding series it was found that some subjects had a strong bias toward or away from the quadrant which had been chosen arbitrarily as the 'correct' one for them in advance of any experimental work with them. This resulted in some cases in too great a number of 'rights' at the beginning of the hour and in other cases resulted in too small a number of 'rights' to permit learning. To eliminate the effect of such bias the 'correct' quadrant for each subject was selected on the basis of his performance in a preliminary exercise of 40 trials where he was told merely to make the dot directly above the tack upon which he had his finger. In this fore-exercise the subject was not told whether his responses were right or wrong; in fact, he was told nothing about there being a right or a wrong response until after the completion of the fore-exercise. We selected as the 'right' quadrant that one which accounted for a percentage of the preliminary trials which was closest to twenty-five. It was hoped that a quadrant could be chosen for each subject on the basis of his performance in the preliminary exercise so that he would show approximately 25 percent of 'rights' at the beginning of the regular series. This hope was not realized. Shifts from the fore-exercise to the regular series occurred in some cases and in others the quadrant had to be chosen rather arbitrarily because most of the responses fell in one quadrant even during the fore-exercise.

Two rather similar sets of directions are given on separate sheets. As marked, one set was used for Series 1 to 5, and the other for Series 6 where the fore-exercise was introduced.

In each series 400 trials per hour were run, with a two minute rest period between each hundred. During these rest periods each subject was required to get up and walk around the room.

In Series 1 only 400 trials were run with each subject. In Series 2, 3, 4, 5 and 6 two sittings of 400 trials each, separated by an interval, were held with each subject. In Series 2 this interval was 48 hours; in Series 3, 4, 5 and 6 the interval was 24 hours.

Directions Used in Series 1, 2, 3, 4, and 5

This is a study to determine your motor skill in locating points. In the mirror you can see reflected a number of figures, each with a brass tack in the center. When we begin you are to put the index finger of your left (non-writing) hand on the tack in the center of each figure as I name it.

(Demonstration here to show subject he is to put his finger on the tack rather than on its reflection in the mirror.)

Now with your right hand take this pencil. As soon as you locate with your left hand the tack in the center of the figure which I name, you are to bring your right hand over in one continuous movement and make a dot on this top paper (which you cannot see) directly above the point upon which you have your index finger. The mirror is only for the purpose of getting your hand on the figure which I name. Pay no attention to the reflection in the mirror in making the dot. Remember, you are to make a dot on the top paper directly above the point upon which you have your finger.

This paper at the right is the rest position. You are to return your hand to it each time as soon as you have made a dot.

If you are very close to the right spot I shall say 'Right'; if you are not close enough I shall say 'Wrong.' The idea of the experiment, of course, is to get as close to the correct point as you possibly can, thus getting as many rights and as few wrongs as possible.[5]

(The continuous sweep with the right hand was insisted upon as the experiment progressed. 'One movement' and 'No hunting' were used as reminders when the subject hovered above a point before making a dot.)

[5] We regret the necessity of misinforming subjects even in so trivial a matter. It seemed unavoidable.

Directions Used in Series 6

This is a study to determine your motor skill in locating points. In the mirror you can see reflected a number of figures, each with a brass tack in the center. When we begin you are to put the index finger of your left (non-writing) hand on the tack in the center of each figure as I name it.

(Demonstration here to show subject he is to put his finger on the tack rather than on its reflection in the mirror.)

Now with your right hand take this pencil. As soon as you locate with your left hand the tack in the center of the figure which I name, you are to bring your right hand over in one continuous movement and make a dot on this top paper (which you cannot see) directly above the point upon which you have your index finger. The mirror is only for the purpose of getting your hand on the figure which I name. Pay no attention to the reflection in the mirror in making the dot. Remember, you are to make a dot on the top paper directly above the point upon which you have your finger.

This paper at the right is the rest position. You are to return your hand to it each time as soon as you have made a dot.

(The fore-exercise of 40 trials was then administered. No 'Right' or 'Wrong' was called. Four cycles of the ten figures in their regular sequence constituted the fore-exercise.)

(The subject was requested to get up and walk around after the completion of the preliminary exercise while the operator computed the percentage of dots in each quadrant for the purpose of determining which quadrant was to be called 'Right' for that particular subject. This quadrant was the one which included most nearly 25 percent of the dots made in the 'no aid' fore-exercise.)

The directions remain the same as before except that I shall now say 'Right' if you are very close to the correct point, and I shall say 'Wrong' if you are not close enough. The idea of the experiment, of course, is to get as close to the correct point as you possibly can, thus getting as many rights and as few wrongs as possible.

(The continuous sweep with the right hand was insisted upon as the experiment progressed. 'One movement' and 'No hunting' were used as reminders when the subject hovered above a point before making a dot.)

(Manuscript received November 18, 1932)

CHANGES IN SKIN TEMPERATURE FOLLOWING INTENSE STIMULATION

BY HARRY HELSON AND LUCENA QUANTIUS

Bryn Mawr College

Among the various correlations between emotional stimuli and their responses those involving changes in temperature have received the least attention at the hands of psychologists. This has probably been due to the fact that no quick, reliable method for measuring changes in temperature has been available until comparatively recently. The measurement of surface temperature presents special problems which have been largely solved through the development of several devices, particularly by the thermo-couple potentiometer which enables one to detect changes almost as soon as they occur and to measure them to a fraction of a degree. The present study concerns the changes in skin temperature following intense stimulation, measured by means of a delicate recording instrument to be described below. We shall show that skin temperature changes in response to various types of stimuli, most of them 'intense,' with as much uniformity as is usually found in the methods of expression, if not more than these methods have usually been found to yield. Bodily changes following stimulation are notoriously variable and so we claim no greater reliability for the results reported here than can be claimed for any work done in this field. The amount of agreement found, however, leads us to believe that further work along these lines may be fruitful.[1]

[1] We shall not attempt any historical summary or detailed discussion of the relations between our findings and other bodily changes following stimulation. This has been done by C. W. Darrow in the following: this JOURNAL, 1927, **10**, 197–226; 1929, **12**, 267–300; *Proc. Ninth Int. Cong. Psychol.*, 1929, 136–7; *Psychol. Bull.*, 1929, **26**, 155–6 and 185–201. For summaries of the literature, particularly on the galvanic skin reflex, see C. Landis, The electrical phenomena of the skin, *Psychol. Bull.*, 1929, **26**, 64–119 (with DeWick), Psychology and the psychogalvanic reflex, *Psychol. Rev.*, 1930, **37**, 381–98. No one, so far as we are aware, has reported as great changes in surface temperature following stimulation as we have found in this study.

It is evident that if a stimulus is effective in causing any change in surface temperature such change will be manifested as either a rise (positive) or fall (negative) in value from some previously existing level or equilibrium. The ideal of research would be to establish a one-to-one correlation between a given stimulus or class of stimuli and either a positive or negative variation in temperature. We might go further and say that ideally opposite changes in temperature should follow opposite stimuli or emotional states, *e.g.*, a change in one direction for pleasant stimuli and one in the opposite direction for unpleasant stimuli, but the history of the methods of expression teaches us that such correlations are at present impossible of attainment. Yet short of this ideal certain uniformities do appear and these we shall now proceed to discuss, after acquainting the reader with the apparatus, stimuli, and mode of procedure.

Apparatus, Procedure, and Stimuli

The temperature recording device was a Leeds & Northrup potentiometer indicator [2] consisting of the sensitive element, an iron-constantan thermo-couple to be held against the skin, galvanometer of the suspension type in a potentiometer circuit, and a balancing dial which gives the temperature in degrees F. when the galvanometer needle gives a null reading. The thermo-couple itself consists of a small wire bent into circular form and having a diameter of about 1/8″. It is not necessary to enter here into the details of the calibration of the unit but calibration against an independent standard showed the instrument we used was 0.5° too high at 86° and 1.15° too high at 103.6°, a fact to be taken into account later. The sensitive element, except for the exposed part held against the skin, is encased in hard rubber tubing containing the lead-in wires to the nstrument. Readings on the dial are given in 0.5° F. and 0.25° can be easily estimated. There is practically no inertia in the needle-indicator which moves the moment there is any change in skin temperature but it takes a few seconds to balance the galvanometer before the amount of change can be determined. So long as the skin is changing in temperature, of course no balance is possible.

A total of 34 *Ss* were used in the experiments, but the first 5 *Ss* were used for preliminary work and will not be considered in this study. 8 of the 29 *Ss* were males. The *Ss* fell into three types: (1) inexperienced both with regard to psychological experimentation and the problem in question; (2) experienced in psychological experimentation but unacquainted with this problem; (3) both experienced in psychology and acquainted with the nature of the problem. In general, while naive *Ss* are best, good

[2] We are indebted to Dr. T. H. Weisenburg of Philadelphia for the use of the measuring instrument and for his interest in the problem. Dr. Weisenburg has been engaged in a study of surface temperature in various neurological conditions. This study was undertaken to determine 'normal' reactions even though the conditions may not have been quite normal.

results were obtained with practically all, the non-reactive *Ss* being experienced as well as inexperienced.[3]

Two sets of temperatures were obtained for each *S*: one in the experimental situation wherein they were subjected to various forms of stimulation described below, and another under conditions as free from stimulation as possible. The experiments were conducted in a windowless room of the laboratory and the room temperature was controlled by means of a thermostat regulating an air-draft system. Usually the temperature of the room was lower at the end of the period than at the beginning, so rising room temperature does not account for the general tendency of skin temperatures to rise under both control and experimental conditions, a fact we shall discuss later.

After preliminary experimentation we decided to use a place below the infra-orbital region near the nose from which to take temperature readings. This region was chosen because of its rich blood supply which comes to the surface easily, its relatively high surface temperature, and its amenability to change. Other parts of the body may be better to work on and may yield different results. Each *S* was given from 10 to 15 minutes in which to get into temperature equilibrium with the room in which we were working but this period, as evidenced by the tendency of the temperature to rise, we now know was too short. Complete equilibrium, we are informed by Dr. H. C. Bazett, may take several hours. Since we are interested in drops as well as rises in temperature, and relative rather than absolute changes, this does not constitute a serious objection to our procedure.

Readings were taken as soon as a balance could be obtained after stimulation and every 30 seconds thereafter, 3 minutes being allowed usually between stimuli for the establishment of level. In the case of the stimuli anticipation of shock, shock, and second anticipation of shock, only 30 seconds were allowed between stimuli for obvious reasons. In some cases even 3 minutes did not suffice for establishment of a new equilibrium or level and in these cases, stimulation coming during a change in temperature might reverse the direction of change or accelerate it, both equally significant from our point of view.

Several possible sources of error should be mentioned lest the reader believe them to have been responsible for some of our findings. Could slight changes in position, pressure, or manner of holding the sensitive element against the face have caused changes in the indicator? These factors were subjected to deliberate test and it was found that slight changes did not affect the galvanometer needle. Changes great enough to cause movement of the needle could be observed by the *E* who sat directly in front of the *S*. In the case of the cold water stimulus thrown on the face, head, or neck it might be objected that the vaso-constrictive effect on one side might be carried over to the side of the face being measured, giving a negative change. While there were more negative changes than positive for this stimulus, nevertheless two positive changes and one zero effect out of a total of 8 shows that this does not necessarily happen. Finally, it may be thought that during sudden, intense stimulation the *S* might remove the thermocouple from the spot or shift it to a new place. Change to a new spot caused little fluctuation in temperature while removal from the face would be apparent to the *E* both from the indicator and directly. In the few cases where this occurred the *S* was informed and replaced it within a very few seconds, whereupon the temperature was found not very different from the value last recorded.

---

[3] At this point we should mention the coöperativeness and good humor of our *Ss* whom we now thank for their part in the experiments. They obeyed our injunction to refrain from discussion of their experiences, took sometimes vicious stimuli with patience, and entered into the spirit of the work. If the reader believes it easy to take a hard, sudden slap on the face without hitting back, let him try it!

The exact stimuli used and manner of presentation merit detailed description in order that future investigations may be made comparable with this one if desired. The experimenter was aided by an assistant who managed the stimuli at a signal from the *E*, the latter tending to the measuring unit. The assistant wore soft slippers and was careful to use every means to prevent the *S* from knowing what was coming next, the *S* being blindfolded.[4] Following is a list of the stimuli and the way in which they were presented:

1. For the stimulus 'anticipation of shock,' *E* said to *S*: "Do you mind a shock?" After the reply of the *S* 30 seconds were allowed to elapse and *E* said: "Now I am going to give it," at the same time closing a circuit which administered a sharp shock to the *S*'s hand. The second 'anticipation of shock' and the giving of the second shock followed in a similar manner, the times being indicated in Figs. 1–4. After this group of stimuli, 3 minutes were allowed during which no stimuli were given and temperature readings taken every 30 seconds.

2. The word-associations group consisted of a neutral word followed by a word or words supposedly having an emotional context for the *S* in question. One to 1.5 minutes were allowed between stimulus words.

3. Ice stimulation consisted in applying a piece of ice to the neck for 30 seconds.

4. The cold water stimulus generally consisted of throwing water on the head, face or neck of the *S*.

5. For intense sound stimulation a hammer was pounded upon a stout oak table approximately two feet from the *S*. The pounding continued for two minutes. It need hardly be said that the pounding was vigorous and not a mild-mannered procedure! Many of the *Ss* reported this as intensely disagreeable.

6. 'Loud noise and scream' consisted in throwing the hammer on the floor, which was tile, and at the same time uttering a blood-curdling (?) scream. This came after a three-minute interval of silence.

7. The hair of some *Ss* was pulled violently.

8. 'Anticipation of pin prick' consisted in saying to the *S*: "Do you mind if I prick you with a pin?"

9. Self-aroused associations were elicited by asking the *S* to introspect upon some pleasant or unpleasant episode with the intention of relating it to *E*. Two minutes later *S* was informed he need not relate it.

10. Odorous stimuli were applied to the nose of the *S* for one minute.

11. Chalk was scraped on the blackboard for one minute.

12. Unpleasant cutaneous stimulation or pain consisted in jabbing the *S* with a needle without warning.

13. The musical stimulus consisted of a record played on a phonograph.

14. The music was suddenly broken off.

15. The music was resumed after an interval.

16. The face of the *S* was slapped during the playing of the record. This was our most intense stimulus and in some cases caused tears.

17. The affectionate stimulation consisted of a kiss.

Not all of the stimuli were given to any one *S*. From 6 to 10 groups of stimuli were given to each *S*, hence not all stimuli were used an equal number of times.

---

[4] Miss Virginia Balough acted as assistant *E*, helped with the making of graphs, and has otherwise earned mention in this study. We are glad to acknowledge her help.

Discussion of Results

The curves in Figs. 1–4 show the changes in temperature of two of the most reactive *Ss* in our group and of two of the least reactive, during the control and experimental or stimulated periods. Stimuli and time are indicated along the abscissa and amount and direction of change along the ordinate.

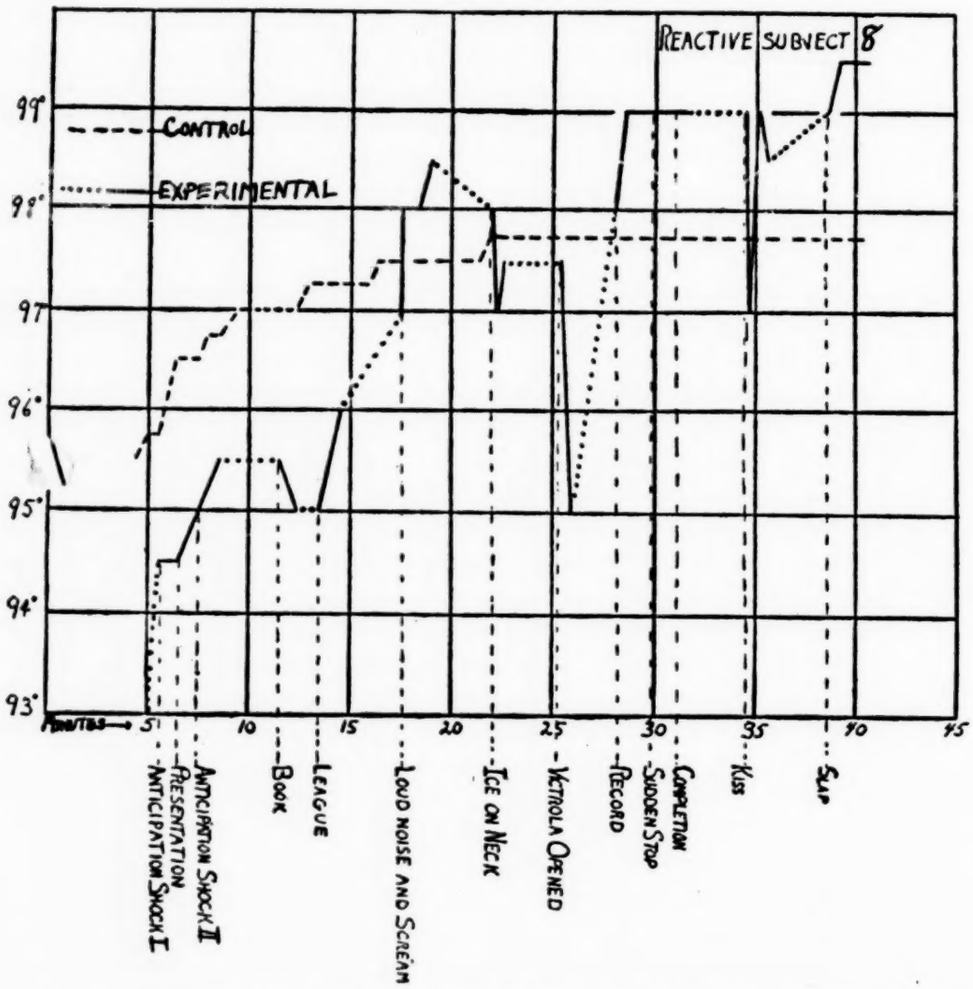


Fig. 1. Showing the changes in temperature under control and experimental (stimulated) conditions for a reactive *S*. Note the absence of negative changes in the control curve.

The curves thus show the type of change caused by each stimulus, the time during which there was zero, positive, or negative change, and the amount of change. That the stimuli were responsible for significant changes in temperature during the experimental periods is borne out by these curves and the data presented in Tables 1–3. The experimental curves

show a larger number of changes, more abrupt changes, and more positive and negative changes than appear in the smoother course of the control curves. The changes in direction of the curves, occurring when a stimulus came before equilibrium had been established, are especially significant. As a rule there is a brief latency between time of stimulation and the response of the indicator but in some cases the indica-

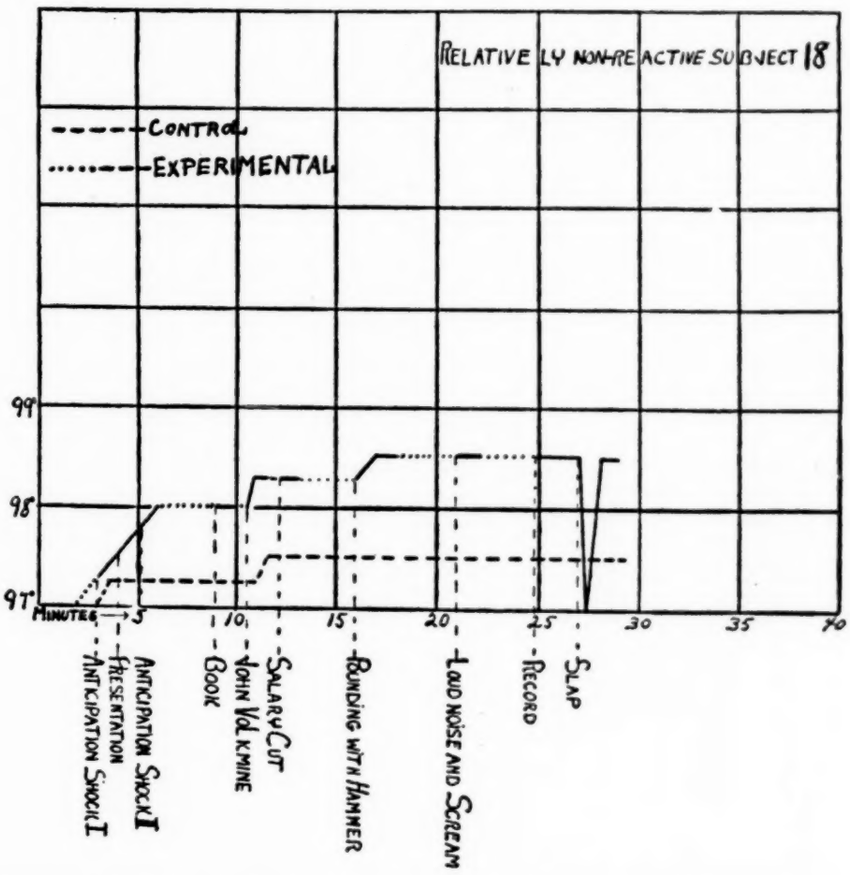


FIG. 2. Showing the changes in temperature under control and experimental conditions for a relatively non-reactive *S*. Both of these curves are more regular than either of the curves in Fig. 1.

tor responded so quickly and decidedly that the temperature change has been drawn from the moment of stimulation.[5] Often the change persisted for more than the interval usually allowed before the next stimulation and then it will be seen that the succeeding stimulus comes during a time of change.

[5] Owing to the fact that we were not measuring the latency of the temperature responses the curves in Figs. 1–4 show changes beginning at the moment of stimulation because the latency is not known.

The fact of individual differences stands out in all of our results. *Ss* differ in their reactions both to specific stimuli and to the stimuli in general, *i.e.*, some *Ss* are more reactive than others both in the control and in the experimental periods. Moreover, a stimulus may evoke a positive response in some *Ss*, a negative response in others, and no response in still

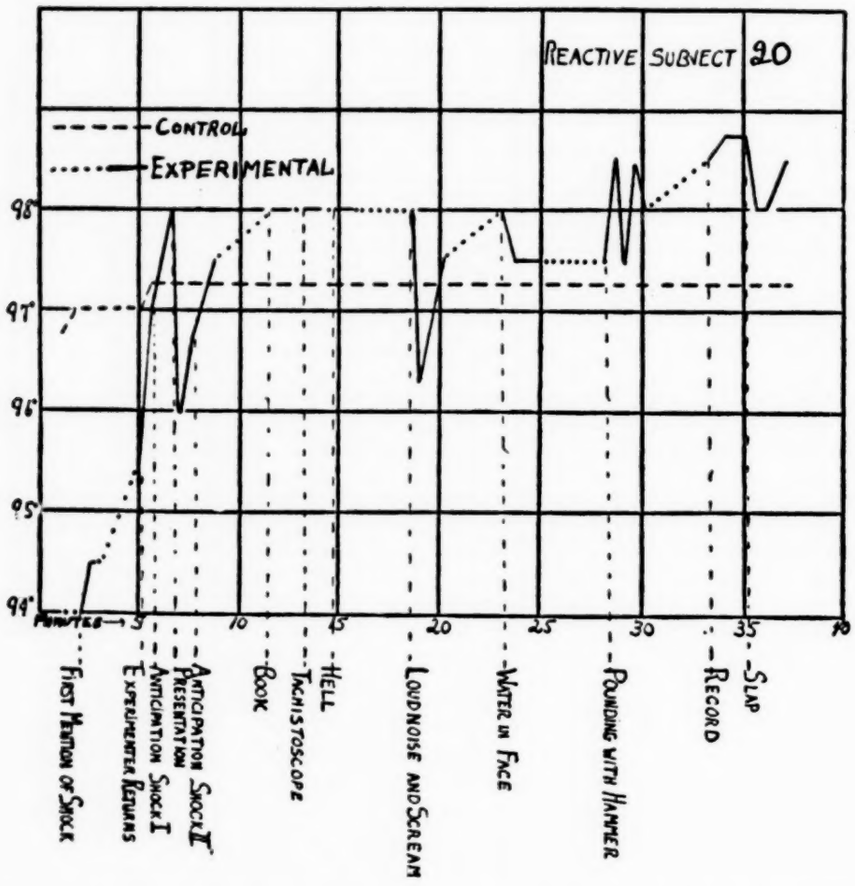


FIG. 3. Showing the changes in temperature under control and experimental conditions for a reactive *S*. Note the almost perfect maintenance of equilibrium under control conditions as contrasted with the sharp changes when stimulated.

others. Thus the slap caused a decrease in temperature in 3 *Ss* and an increase in temperature in one of the 4 *Ss* pictured in Figs. 1–4. Nevertheless some stimuli show a preponderance of positive or negative changes, a fact which can be seen from the data presented in Table 2. In general we found that the more reactive *Ss* in the experimental conditions were more apt to show changes during the control period when no stimulation was used. Analysis of the data fails to reveal signifi-

cant differences between the men and women, between the experienced and inexperienced *Ss*.

In Table 1 we have presented the total number, sum and average of the absolute changes in temperature for each of the 29 *Ss* under control and experimental conditions. Whereas the average change in the control period was 0.273°, that for

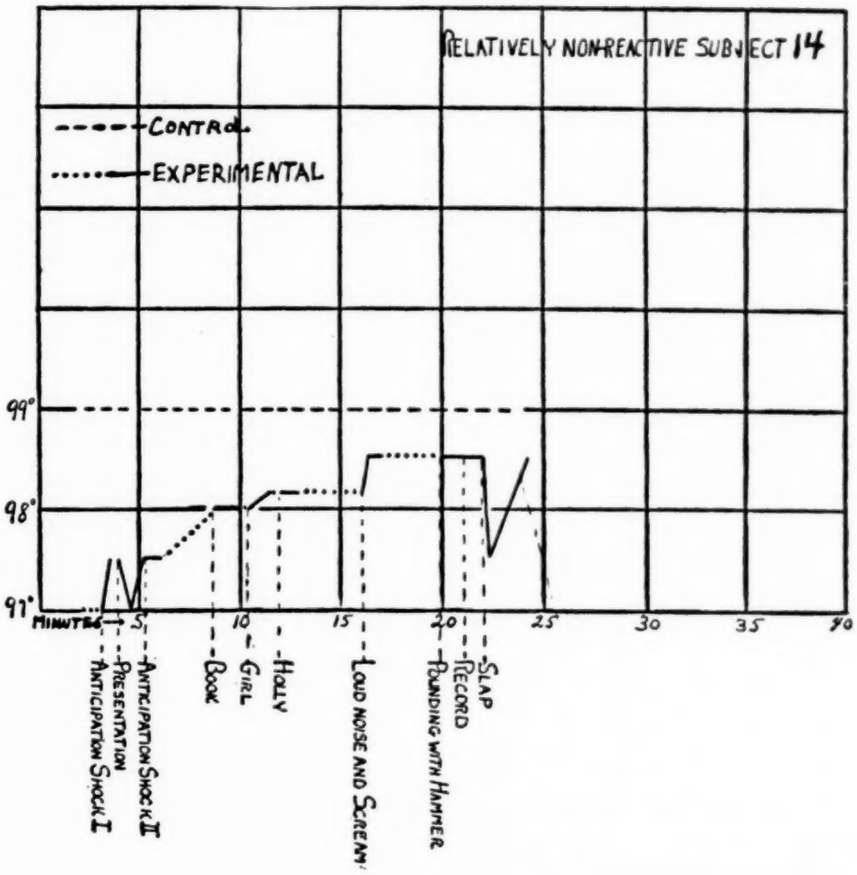


FIG. 4. Showing the changes in temperature under control and experimental conditions for a relatively non-reactive *S*. The control curve is the most constant in our records due to the fact that the *S* happened to be in equilibrium with the temperature of room during the control measurements.

the experimental series was 0.533°. The difference between the total number of changes and the average amount of change for the two periods is statistically reliable, as shown by the values given in the table. Taking individual results we find only one *S*, 19, who shows a larger number of variations during the control than during the experimental period, but even in the case of this *S* it must be noted that the average amount of

TABLE 1

SHOWING CHANGES IN SURFACE TEMPERATURE DURING CONTROL CONDITIONS AND UNDER EXPERIMENTAL CONDITIONS

Control Conditions				Experimental Conditions			
Subject	Number of Changes	Sum of Changes	Average of Changes	Subject	Number of Changes	Sum of Changes	Average of Changes
5	2	.50	.25	5	14	8.50	.61
6	3	1.00	.33	6	20	10.25	.51
7	10	3.00	.30	7	16	9.25	.58
8	8	2.25	.28	8	20	15.50	.78
9	5	2.00	.40	9	10	6.50	.65
10	8	2.00	.25	10	18	7.75	.43
11	12	3.50	.29	11	14	7.50	.54
13	8	2.00	.25	13	13	7.50	.58
14	0	0.00	.00	14	10	4.50	.45
15	8	2.50	.31	15	15	6.50	.43
16	3	1.00	.33	16	7	3.75	.53
17	8	2.75	.34	17	12	8.50	.71
18	2	.50	.25	18	7	4.25	.61
19	13	3.25	.25	19	10	5.50	.55
20	2	.50	.25	20	21	17.50	.83
21	6	2.50	.42	21	11	7.00	.64
22	4	1.00	.25	22	10	4.50	.45
23	4	1.00	.25	23	15	7.50	.30
24	2	.50	.25	24	14	6.25	.45
25	6	1.50	.25	25	9	3.75	.42
26	6	1.50	.25	26	14	6.50	.46
27	8	2.50	.31	27	16	9.00	.56
28	4	1.00	.25	28	13	5.50	.42
29	10	3.00	.30	29	13	6.75	.52
30	2	.50	.25	30	15	5.25	.35
31	12	3.25	.27	31	14	10.25	.73
32	9	2.50	.28	32	19	10.50	.55
33	3	.75	.25	33	5	2.00	.40
34	9	2.25	.25	34	27	12.00	.44

Total changes = 177 Av. = .2728 Total changes = 402 Av. = .5334
Av. no. changes = 6.10 $\sigma_{dist.}$ = .075 Av. no. changes = 13.86 $\sigma_{dist.}$ = .1245
$\sigma_{av.}$ = .0139 $\sigma_{av.}$ = .0231

$\sigma_{diff.}$ = .0264

$$\frac{\text{Diff. of averages}}{\sigma_{diff.}} = 9.868$$

change was greater in the experimental period than in the control period, being 0.55° as against 0.25° for the non-stimulated conditions. Our conclusions hold therefore not only for the group as such but also for each individual, in spite of the fact that some *Ss* are relatively non-reactive.

Questions as to the effectiveness, maximum change, number of positive, negative and zero changes, and average change

produced by each stimulus are answered in Table 2. The stimuli in this table have been ordered, with one exception, according to the average change caused by each. The kiss stands first with the highest average change, followed by the slap, water thrown on face or neck, hair pulling, and hammer pounding as next in order of effectiveness. The stimuli causing the largest variation in temperature are water thrown

TABLE 2

SHOWING NUMBER OF POSITIVE, NEGATIVE, AND ZERO CHANGES, TOTAL NUMBER OF CHANGES, THE MAXIMUM CHANGE AND THE AVERAGE CHANGE FOR EACH STIMULUS

Stimulus	Positive Changes	Negative Changes	Zero Changes	Total Number	Maximum Change °	Average Change °
Kiss	1	5	0	6	2.00	1.92
Slap	5	22	1	28	2.50	1.65
Water thrown on face or neck	2	5	1	8	4.00	1.38
Hair pulled	3	8	0	11	2.00	0.93
Hammer pounding	10	1	5	16	1.25	0.69
Word association	17	2	5	24	1.00	0.65
First anticipation of shock	23	1	4	28	1.50	0.52
Shock given	14	8	5	27	2.00	0.52
Second anticipation of shock	23	0	4	27	1.50	0.50
Loud noise and scream	13	5	9	27	1.75	0.46
Ice on neck	1	2	3	6	1.00	0.38
Anticipation of pin prick	4	1	0	5	0.50	0.35
Playing of record	12	4	12	28	1.00	0.29
Music stopped suddenly	1	1	2	4	0.50	0.31
Music completed	1	0	3	4	0.50	0.25
Self-aroused associations	1	1	2	4	0.50	0.25
Pleasant odor	1	0	3	4	0.50	0.13
Chalk scraping	1	0	1	2	0.50	0.13
Pin prick	0	0	3	3	0.00	0.00

on face or neck with a maximum of 4°, the slap with a maximum of 2.50°, and the kiss and hair pulling with a maximum for each of 2.00°. It is interesting to note that the first anticipation of shock gave as large variation, in the average, as the actual shock. Every one of the stimuli used was effective with one or more of the *Ss* with the exception of the pin prick, which might have been effective had it been tried with more *Ss*. Since we were concerned with establishing the fact of temperature change following stimulation and were desirous of finding effective stimuli, stimuli which were relatively ineffective were not used as much as more effective ones. However, the converse is not true: the most effective stimuli were

not used as many times as some less effective, so that while our results may have been favored by a tendency to use poor stimuli less, they were not favored too much by too great use of some of the better stimuli.

It will be seen that the first 4 stimuli tend to cause drops in temperature *at the spot measured* whereas the next 6 stimuli tend to cause a rise in temperature. The median change occurs at the loud noise and scream, the range from median to most effective being from 0.46° to 1.92°. Below the median stimulus there seems to be no clear-cut effect upon temperature change as the numbers of positive, negative and zero changes tend to be nearly equal, showing a more or less random distribution. Opposite stimuli do not seem to cause opposite changes, although our stimuli were mostly unpleasant, outnumbering the pleasant ones. Even in the case of the kiss we cannot be sure that it was pleasant—in fact one of the *Ss* (male) reported it very unpleasant 'in these circumstances.' The number of increases in temperature are much greater than the negative changes, the total number being 133 positive, 66 negative and 63 instances of no change. Experiments with either our own or other *Ss* might well change the order of effectiveness, predominant type of change caused by each, or the value of any or all of the stimuli. Furthermore, we cannot be sure that a given type of change will be found all over the body for a given stimulus: a drop or rise in temperature in one part may be accompanied by a rise or drop in temperature in another part of the body.

We have used the term 'intense stimulation' in order to avoid the use of terms which imply internal or emotional states because we have found, in line with the generally accepted fact, that bodily changes do not always mean corresponding introspective changes. To be sure, large or sharp variations in the temperature curves were often found to accompany marked changes in feelings or emotions, but occasionally this was not true. Stimuli may cause internal effects which do not appear on our records, or they may be reported as ineffective when the bodily change is noticeable. Further work is needed to establish the exact correlation which may

be assumed to exist between temperature changes and introspectable emotional states.

Inspection of Figs. 1–4 reveals that skin temperature tends to rise in both the control and experimental periods, in spite of the frequent drops which occur during the latter. In our preliminary experiments where the tip of the finger was used for measurements we noted a tendency for the skin temperature to *drop* as time went on. Does the rise found in the face mean that blood is being withdrawn from other parts of the body to the part being measured on account of the irritation caused by the iron-constantan couple? Before attempting to answer this question let us look at Table 3. In it will be found the average temperature of all *Ss* at the beginning and end of the control and the experimental periods. There is a difference of 1.33° for the control and 2.42° for the experimental period. Calculation of the significance of the difference shows that the temperature is significantly higher at the end of each period than it was at the beginning. Moreover, the increase at the end of the experimental period is significantly higher than that of the control period, as shown by the ratio 3.37 of the difference between the averages of the differences divided by its standard deviation. In spite of the fact, therefore, that the temperature rises in both periods, we know that the rise under experimental conditions is significantly greater than without stimulation. This is but a check on the facts presented in Table 2 wherein we found more increases than decreases in temperature following stimulation.

What caused the temperature to rise under normal, control conditions, for we have found this rise significant in itself? The following seem to be the most plausible explanations, capable of further check: (1) the *Ss* may not have been given enough time to reach an equilibrium with their surroundings, some of them having come in from the outside; as stated above, hours may be required for this purpose; (2) the cooling effect of the room in which the *Ss* were measured may have been less than that of the surroundings in which they were previously, regardless of the absolute temperature, since cooling depends upon other factors as well; (3) the pressure of the wire on the

face may have set up a local irritation which caused the skin underneath to undergo an increase in temperature; (4) the thermo-couple sank into the skin in many cases and may have caused an approximation to internal temperature; (5) the *Ss* were under more or less constant tension during both the experimental and control periods, some believing that something was going to happen even during the latter conditions.

We now approach the question of the abnormally high surface temperatures recorded in some cases, indicating an abnormally high internal temperature. We found it impossible to make measurements simultaneously of external and internal temperatures, although it may be done fairly easily under some conditions, and so cannot say what happened to the internal temperatures of our *Ss*. While the temperature of the part we measured is generally higher than that of other parts of the body which are not covered, it still should be several degrees below the internal temperature. Our instrument recorded temperatures as high as 101°, a value which would make the internal temperatures very high indeed. One source of the high readings may have been due to the instrument itself since we found that it gave readings somewhat high. But this will account for only a degree at the most. Possibly the irritation of the wire on the face may have raised the temperature above normal, or we may have approached internal temperatures when the thermo-couple sank deeply into the skin. In any case the few high temperatures which we found do not invalidate our results since we were interested in changes from a pre-existing level or in accelerated changes. Enough drops in temperature were found at all levels to show that our stimuli were truly effective in lowering as well as in raising the temperature.

Interpretations and Conclusions

In this preliminary study it is not our intention to enter into any long discussion of the causes of the changes in temperature we have found or their relations to the many other known responses to intense stimulation.[6] It is very probable

[6] See note 1, *supra*.

TABLE 3

GIVING THE AVERAGE TEMPERATURE AT THE BEGINNING AND END OF THE CONTROL AND EXPERIMENTAL PERIODS WITH THE STATISTICAL VALUES NECESSARY FOR DETERMINING THE RELIABILITY OF DIFFERENCES BETWEEN THE TWO

	Control Period	Experimental Period
Temp. at beginning	96.26 ± 0.96	96.00 ± 1.76
Temp. at end	97.59 ± 1.09	98.42 ± 1.10
Difference	1.33	2.42
$\sigma_{Diff.}$	0.27	0.39
Diff./$\sigma_{Diff.}$	4.92	6.24
Average difference	1.33 ± 0.88	2.41 ± 1.48
$\sigma_{Av.\ Diff.}$	0.16	0.25
Difference between the average differences	1.08	
$\sigma_{Diff.}$	0.32	
Diff. between the average diff./$\sigma_{Diff.}$	3.37	

that the chief source of our changes is to be sought in the vaso-motor system controlling the blood supply to the part measured.[7] Vaso-constriction and vaso-dilatation may occur through autonomic innervation resulting in decreased or increased capillary circulation with corresponding changes in temperature. Some of the stimuli may have influenced the physiological oxidations to cause an actual change in heat production throughout the entire body but of this we cannot be sure. It may explain the gradual rise in temperature. Since our stimuli probably affected the secretory mechanisms which are known to play a part in heat loss through the evaporation of sweat from the skin, surface temperature may have been affected through this avenue as well. It is believed that the most important factor in heat control is exercised through the sweat and vasomotor nerves by reflex action through the heat centers, if such exist. It should be remembered that vaso-dilatation may cause a lowering of general body temperature by resulting in a greater supply of blood to the skin where it is lost by conduction and radiation to the cooler air. But the temperature of the skin as such may be increased through vaso-dilatation if the amount of heat lost is less than the amount produced in this manner.

We should expect intense stimulation to have autonomic effects and it would follow that, among other effects produced

[7] For a general discussion of temperature control, see *A Textbook of Physiology*, by W. H. Howell, 1919, 7th ed., pp. 967–978.

by autonomic action, temperature changes which are regulated by autonomic fibers will also occur. But, as Darrow has pointed out,[8] the action of the vaso-constrictors and vaso-dilators is by no means simple and we have no real insight into the true physiological mechanisms underlying temperature and other autonomic functions closely associated with it. The lack of knowledge regarding the exact mechanisms involved in these changes need not prevent us from proceeding with our attempted correlations between stimuli and responses.

Summarizing our findings we may say that:

(1) Changes in skin temperature occur in response to intense stimulation and are as easily demonstrable as the galvanic skin reflex and other bodily correlates of emotional states usually investigated in the so-called methods of expression.

(2) Variations in temperature as large as 4° have been found and the changes may be either positive or negative in character. The positive changes in these experiments outnumbered the negative in a ratio of nearly 2 : 1.

(3) Individual differences due to differences in response to particular stimuli and general reactivity prevent us from generalizing as to the effect of specific stimuli, although there is a tendency for some stimuli to cause predominantly positive changes while others show more negative changes.

(4) 26 out of the 29 *Ss* showed responses of at least one degree to some one or more of the stimuli, while the remaining three *Ss* reacted with at least 0.5° to some stimulus, except the pin prick.

(5) Almost any sort of stimulation will cause temperature change if intense and sudden enough, or if it possesses 'stimulating character' such as is found in crucial words of the word-associations test. Often anticipation of an unpleasant stimulus caused more change than the stimulus itself.

(6) The presence or absence of emotional states cannot be inferred from the temperature response although there is a tendency for large changes to be accompanied by reportable internal changes.

[8] *Opp. cit.*

We have tried to indicate various sources of error, answer objections and suggest further lines of research. Many interesting questions regarding the effect of intense stimulation on temperature responses still remain to be answered.

(Manuscript received December 1, 1933)

THE EFFECT OF VARYING EXTERNAL CONDITIONS ON LEARNING, RETENTION, AND REPRODUCTION [1]

BY A. H. MASLOW

University of Wisconsin

This paper is concerned with the effect of varying external conditions on learning, retention, and recall. Since Wilson's first experiment on this problem, there have been published only three experimental studies that attack the problem directly. Their bearings on the question of economy in learning and educational psychology in general are many and they afford as well an experimental approach to the testing of Gestalt principles of learning.

The question may be put as follows. Is it advantageous to have external conditions as constant as possible during the learning, retention, and reproduction process? Which is least to be desired, change during the learning period, during retention, or during recall? Just how specific is learning? If it takes place against a varied external background, will the removal of any outstanding part of this background cause disturbance? That is, does the material being learned become part of a gestalt composed of an integrated room-chair-desk-learning situation or will this material be independent of the background against which it is learned?

Historical Review—Observations on Animals.—These experiments were uniformly concerned with the maze behavior of several animals; the rat, the guinea pig, pigeon, etc.

Small (**13**), Carr (**3**), and Watson (**15**) with the white rat, and Allen (**1**) with guinea pigs found that changing the environment in various ways, after the maze had been learned, in no way decreased the efficiency of this habit. Yerkes (**16**) with the frog, and Rouse (**12**) with the pigeon found that such change did affect the maze running habit. It should be

[1] I wish to thank Professor Hulsey Cason for suggestions and criticisms in the course of the experiment.

pointed out that both experiments varied the visual cues of the animal, Yerkes changing the coloring of the maze wall and Rouse removing the light after the maze had been learned. We are concerned in this experiment with the effect of changing background conditions at the same time that the cues actually used in learning are not affected. While varying the visual cues in the case of the rat did not affect his learning, since they are not essential after learning has occurred, these visual cues may be essential for the pigeon and the frog.

One interesting contradiction to the experiments cited above was that observed by Watson (**15**). Rotating the maze in respect to compass direction had a marked and unexpected disturbing effect on maze running performance. This finding was confirmed by Porter (**11**) working with sparrows, by Hunter (**7**) with pigeons, and by Dashiell (**5**) working with rats. This latter experimenter attributed this phenomenon to disturbed bodily orientation.

Observation on Humans.—Smith and Guthrie (**14**) report an unpublished experiment by Wilson. He had two groups of subjects learn in a laboratory room. He then tested one group for recall in the same room, and the other group in the open air. Two more groups learned in the open air, one being tested in the same surroundings, the other in the laboratory room. He found that the two groups which learned and recalled in the same surroundings (whether room or open air) did better than the groups that learned in one place and were tested in another. He repeated the experiment, using the odor of peppermint as the varying condition and obtained similar results.

Shuh Pan (**10**) studied the dependence of learning and recall upon the context in which the material appeared. This context was composed of pictures, words, and numbers (Related and unrelated to the material). His results were various and interesting. A word context, logically related to the response word, exerted a beneficial effect on learning. Its removal during recall was highly detrimental. When the context was varied during the learning, the detrimental effects were lessened. A word context, logically unrelated to

both stimulus and response words, was detrimental to learning. Its removal during recall was beneficial. The presence of a number context in learning was slightly detrimental and its removal in recall exerted a similar effect.

Bills (2) found that a condition of tension during learning helped this learning. Tension during recall was also helpful.

Pessin (17) set himself the same problem. If an individual learns material under certain laboratory conditions, will he be able to relearn this material more efficiently under similar or under dissimilar conditions? The two conditions used were (*A*) a combined auditory and visual stimulation (a bell and bright light) and (*B*) a condition of relative quiet. Four combinations of *A* and *B* were possible:

A for learning; *A* for relearning
B for learning; *B* for relearning
A for learning; *B* for relearning
B for learning; *A* for relearning.

All four were tried. A series of nonsense syllables was used for learning and relearning after four minutes. The results were inconclusive. No significant differences were found for any of the four sets of conditions.

Reed (18) has contributed a most significant paper on the problem. In a series of three experiments, she tried (I) varying conditions by changing the position of the subject, half of the subjects standing and half sitting during learning, and half of each of these groups sitting and half standing during recall. Thus there were four groups, as in Pessin's experiment, one learning and recalling while sitting, one while standing, a third group learning while sitting and recalling while standing, and finally a group learning in a standing position and recalling while seated. (II) Varying the sensory modes used in learning and recalling (visual and auditory). (III) Varying the form of response used (oral response and written response).

She found that in Experiments I and III, similar conditions for both learning and recall were not better than changed conditions. She did, however, get a significant

difference in Experiment II, that is, changing the sensory mode of presenting stimulus words in the recall test caused the recall scores to be lowered and the number of good scores to be decreased.

These results support both the Gestalt theory (which insists on the importance of the total stimulating situation in determining the response) and the simple-machine theory (which speaks rather of point-to-point connections in the sensory-motor system). The kinæsthetic cues may be changed and not affect the learning arc. The same holds true for the form of response to the stimulus words. If the sense modality used for presenting the stimulus words is changed, the efficiency of recall will be affected.

Procedure and Apparatus.—These experiments were carried on at the University of Wisconsin in 1930 and 1931, the subjects being the writer's students in the Psychology I course. For experiment No. 1, $n = 22$; for experiment No. 2, $n = 35$. The writer served as experimenter.

The material used for learning consisted of 100 cards. On each card were typewritten nine meaningful, unassociated, three-letter words. The words were picked by chance from the list of 'Familiar Words for Learning Experiments' in Cason's article on backward association (4). This list was compiled by him from a list of the most common three-letter words in Thorndike's '*Teachers' Word Book*' (1921). Although these words were chosen by chance, certain precautions were taken. All pairs of words with obvious associations were broken up. Words beginning with the same letter or rhyming with each other were never put near each other. Meaningful groups of words were split up and every attempt was made to have the series equal to each other in memory value.

A table top partition (26″ × 48″), painted dead black, was placed down the long axis of the table between the experimenter and the subject. Exactly in the center of this partition was a slit-like window (1″ × 5″) framing the card which the subject was to see. From his side the experimenter was able to control a simple slide mechanism similar to that used in moving and replacing slides in a projection lantern.

Each card was exposed for ten seconds, during which time the subject had just enough time to read the series twice. Thus the learning was of the memory span type. At the end of the ten seconds, the card was withdrawn and its place was taken by a blank white card at which the subject gazed during the period of retention (5 seconds in Experiment I, 15 seconds in Experiment II). At the end of the retention period, the experimenter said 'Ready' and the subject repeated all the words he could remember out of the series of nine. Only perfect reproduction of the words was scored correct. As many series were shown as was possible in an hour. At the end of a half hour the subject rested for three minutes. To avoid practice effects, the first eight series was not counted. It was noticed that after the rest period the subject usually got an abnormally high score on the first card shown. This was also discarded.

The subject was instructed to read through the list twice and told not to pay any attention to the bell. Beyond this he was given no hint as to the purpose of the experiment.

In Experiment I the pause between cards varied from five to fifteen seconds. It was noticed that the sooner one series followed another, the poorer were the results. In Experiment II the variable was controlled, exactly 10 seconds elapsing between the end of the reproduction period and the presentation of the next card. This interesting phenomenon became the subject for another experiment which will be reported later.

The sound of a bell was chosen as the variable external factor because of its simplicity and convenience. The two conditions used were, therefore, (1) bell and (2) silence. All possible combinations of the three periods and the two conditions were used. They are as follows:

TABLE 1

Series	Learning	Retention	Reproduction
S	Silence	Silence	Silence
T	Silence	Silence	Bell
U	Silence	Bell	Silence
V	Silence	Bell	Bell
W	Bell	Bell	Bell
X	Bell	Bell	Silence
Y	Bell	Silence	Bell
Z	Bell	Silence	Silence

These series were ordered in such a way as to eliminate the factors of position, serial association, and similarity of series. The word cards were presented so that each was learned under every one of the eight sets of conditions an equal number of times.

Two separate experiments were completed. In the first, the period of learning was 10 seconds, retention 5 seconds, and reproduction 15 seconds. In the second experiment the periods were 10 seconds for learning, 15 for retention, and 10 for reproduction. The principal difference between the two procedures thus lay in the time intervals.

The method used in this study differs from those hitherto used in several respects. The learning process chosen is the simplest one possible, namely, memory span learning. Using this process, a complete act of learning was possible in our experiment, in 30 seconds for Experiment I, and in 35 seconds for Experiment II. We could go through as many as 50 series for each subject. This of course would be impracticable with any other type of learning material. We were able to try more combinations with each subject and repeat them more often than would have been possible had we used, say, the savings method with a long list of nonsense syllables.

The simplicity of the method was of course in some ways a defect. Slight changes in external environment would not upset a simple function as much as they would a more complex one. It was because of this that our critical ratios were not higher than they were. Then too we have no separate measures for learning and recall as we would have had using the savings method.

Results

The results for Experiment I and Experiment II will be presented separately and then compared with each other.

The order of efficiency for the various conditions was as follows (Table 2):

TABLE 2

MEANS FOR EACH OF THE SERIES (IN ORDER OF EFFICIENCY)

	Learning	Retention	Reproduction
W—5.34 ± .08.....	Bell	Bell	Bell
U—5.24 ± .07.....	Silence	Bell	Silence
S—5.22 ± .09.....	Silence	Silence	Silence
T—5.16 ± .08.....	Silence	Silence	Bell
Z—5.10 ± .08.....	Bell	Silence	Silence
V—5.09 ± .08.....	Silence	Bell	Bell
Y—5.08 ± .08.....	Bell	Silence	Bell
X—5.04 ± .08.....	Bell	Bell	Silence

The fact that the W series is the best for learning has no clear cut significance. We might have expected this result from what we know of distraction experiments. However, we are not testing the effect of distraction. On the basis of what we know about the effects of distraction, we might expect that the more bell, the better the learning. But this interpretation is excluded by the fact that the X and Y series are at the bottom. It will be more profitable to compare the results of the three periods. We might add for each period the position numbers, as follows (using the table of order of efficiency in Table 2). For example, if we look at this table we see under the head of 'Learning' the condition used in the various series during the learning period. Do we find bell or silence at the top of the list, that is, in the most efficient series?

	Learning Period		Retention Period		Reproduction Period	
	Bell	Silence	Bell	Silence	Bell	Silence
Position.....	1	2	1	3	1	2
Position.....	5	3	2	4	4	3
Position.....	7	4	6	5	6	5
Position.....	8	6	8	7	7	8
	21	15	17	19	18	18

The only conclusion we can draw from this treatment is that during the learning period, it is more advantageous to learn in silence.

TABLE 3

DIFFERENCE BETWEEN EFFICIENCY OF SERIES

Experiment 1

T	U	V	W	X	Y	Z	
0.055 ± .12	0.026 ± .04	0.126 ± .12	0.123 ± .11	0.176 ± .12	0.136 ± .12	0.120 ± .12	S
	0.081 ± .11	0.070 ± .11	0.179 ± .11	0.120 ± .12	0.80 ± .11	0.064 ± .12	T
		0.151 ± .11	0.098 ± .11	0.201 ± .11	0.161 ± .11	0.145 ± .11	U
			0.249 ± .11	0.050 ± .12	0.010 ± .11	0.006 ± .11	V
				0.299 ± .11	0.259 ± .11	0.243 ± .11	W
					0.040 ± .11	0.056 ± .12	X
						0.016 ± .11	Y

There is one deduction that we might legitimately make from the results of Experiment I. We can say that when the conditions for learning and for reproduction are the same, the learning is more efficient. The retention period is not as important so far as varying conditions go. If we take the series that conform to this condition, namely, that conditions for both learning and recall were the same, their position numbers are in order—1, 2, 3, and 7 or W, U, S, and Y. Why the Y series is so far down is impossible to explain. The position numbers of the series in which conditions for learning and reproduction are different are 4, 5, 6, and 8 (T, Z, V, X).

These results seem to show that varying external conditions *do* lower the efficiency of this learning process.

This view is consistent with Wilson's experiment (**14**), referred to above. Gates (**6**) also believed that similarity of conditions during learning and recall are more conducive to learning. The results of Experiment II do not substantiate the tentative theory, as well as do those of Experiment I, however.

If we examine Table 3 we find that there are no reliable differences between any of the means of the series in Experiment I. The only differences that approach reliability are those between series W (at the top of the list) and the series X, V, Y, and Z at the bottom of the list. These differences have critical ratios of more than two.

Experiment II

In general, this experiment was exactly the same as Experiment I except for changes in the time of the various periods. The retention period now is given 15 seconds, instead of 5 as in Experiment I, and the reproduction period is now 10 seconds instead of 15.

Our results in Experiment II are more reliable than those in Experiment I. The order of efficiency was:

Here again we have the W series at the top. The order of the series (except for interchange of U and Y) is approxi-

TABLE 4

MEANS FOR EACH OF THE SERIES IN EXPERIMENT II (IN ORDER OF EFFICIENCY)

	Learning	Retention	Reproduction
W—5.24 ± .05	Bell	Bell	Bell
T—5.13 ± .07	Silence	Bell	Bell
Y—5.02 ± .07	Bell	Silence	Bell
S—5.01 ± .08	Silence	Silence	Silence
V—4.93 ± .07	Silence	Bell	Bell
Z—4.86 ± .07	Bell	Silence	Silence
X—4.85 ± .08	Bell	Bell	Silence
U—4.83 ± .08	Silence	Bell	Silence

mately the same as the order for Experiment I. If we put them side by side we can see this more readily.

Experiment I	Experiment II
W	W
U	T
S	Y
T	S
Z	V
V	Z
Y	X
X	U

Except for series U and Y, the rank order of all is the same or approximately the same. The explanation for this interchange of position must be found in the increasing period of retention.

Series U and Y are the converse of each other, one being

U =	Silence	Bell	Silence
Y =	Bell	Silence	Bell.

The bell ringing during a long period of retention would undoubtedly explain the relative loss in efficiency in series U, at the same time that it would increase the relative advantage of series Y.

If we examine the support for the tentative generalization made as the result of Experiment I, we find that it does not stand up quite as well as it did then. If we add up rank numbers for series that have similar conditions for learning and recall, we have 1, 3, 4, 8 (W, Y, X, U) as against 2, 5, 6, 7 (T, V, Z, X). Adding them reveals only a slight advantage (16 vs. 20) for the similar series as over against the dis-

similar series. This would naturally be expected with the increased importance of the retention period.

If we get the differential effects of the bell on the three periods we find them to be as follows:

	Learning Period		Retention Period		Reproduction Period	
	Bell	Silence	Bell	Silence	Bell	Silence
Position.....	1	2	1	2	1	4
Position.....	3	4	5	3	2	6
Position.....	6	5	7	4	3	7
Position.....	7	8	8	6	5	8
	17	19	21	15	11	25

We see silence as an advantage during the retention period. Introspective reports of the subjects, after the experiment, confirm this finding. Unexpectedly enough we find that the bell during the reproductive period is a great aid to efficiency, whereas in Experiment I there was absolutely no difference. I have no explanation to offer for this phenomenon.

If we examine Table 5 we find that we have differences much more reliable than Experiment I.

If we arrange the series in the order of their efficiency (W, T, Y, S, V, Z, X, U), we find that we have differences that are reliable or that approach reliability, for any series and any other series, two rank orders removed from it. That is, the difference between S and X, V, and W will be more reliable but the difference between S, Y, and V will not be.

Summary

Using memory span learning, two experiments were carried out to find the effects of varying external background conditions on the efficiency of the learning process. They are interpreted as giving some slight support to the theory that similar conditions for learning and recall are more efficient than dissimilar conditions. The advantages and disadvantages of the method are discussed.

(Manuscript received December 10, 1932)

TABLE 5
DIFFERENCES BETWEEN AVERAGES OF SERIES
Experiment 2 ($n = 35$)

T	U	V	W	X	Y	Z	
0.112 ± .10	0.190 ± .11	0.090 ± .10	0.219 ± .09	0.166 ± .11	0.007 ± .11	0.161 ± .10	S
	0.302 ± .10	0.202 ± .10	0.107 ± .08	0.278 ± .10	0.106 ± .10	0.273 ± .10	T
		0.0996 ± .10	0.408 ± .09	0.024 ± .11	0.196 ± .11	0.029 ± .10	U
			0.309 ± .08	0.075 ± .10	0.097 ± .10	0.071 ± .10	V
				0.384 ± .09	0.212 ± .09	0.380 ± .08	W
					0.172 ± .10	0.004 ± .10	X
						0.168 ± .10	Y

REFERENCES

1. ALLEN, MISS, The associative processes of the guinea pig, *Jour. Comp. Neurol. and Psychol.*, 1904, **14**, 293–359.
2. BILLS, A. G., Influence of muscular tension on efficiency work, *Amer. J. Psychol.*, 1927, **38**, 227.
3. CARR, HARVEY, Maze study with white rats. I. Normal animals, *J. Animal Behav.*, 1917, **7**, 257.
4. CASON, HULSEY, Specific serial learning, A study of backward associations, J. EXPER. PSYCHOL., 1926, **9**, p. 202.
5. DASHIELL, J. F., Direction orientation in maze running by white rats, *Comp. Psychol. Monog.*, 1930, **7**, No. 2.
6. GATES, A. I., Recitation as a factor in learning, *Arch. of Psychol.*, 1917, **6**, No. 40.
7. HUNTER, W. S., Some labyrinth habits of domestic pigeons, *J. Anim. Behav.*, 1911, **1**, 278–304.
8. LASHLEY, K. S., *Brain Mechanisms and Intelligence*, 1929, Chicago, p. 137.
9. LASHLEY AND HUBBERT, Retroactive association and elimination of errors in the maze, *J. Anim. Behav.*, 1911, **1**, 130–138.
10. PAN, S., The influence of context on learning and recall. J. EXPER. PSYCHOL., 1926, **9**, p. 468.
11. PORTER, J. P., Further study of English sparrow and other birds, *Amer. J. of Psychol.*, 1906, **17**, 248–271.
12. ROUSE, E. A., The mental life of the domestic pigeon, *Harvard Psychol. Stud.*, 1906, **II**, 581–613.
13. SMALL, W. S., Experimental study of the mental process of the rat, II, *Amer. J. of Psychol.*, 1900, **12**, 232. (Quoted by Watson.)
14. SMITH, S., AND GUTHRIE, E. R., *General psychology in terms of behavior*, 1921. New York, p. 112.
15. WATSON, J. B., Kinæsthetic and organic sensations; their role of reaction of white rats in the maze, *Psychol. Mongr.*, 1907, **8**, pp. 33.
16. YERKES, R. M., *Harvard Psychological Studies*, 1903, **1**, p. 579.
17. PESSIN, J., Effect of similar and dissimilar conditions upon learning and relearning, J. EXPER. PSYCHOL., 1932, **15**, p. 427.
18. REED, H., The influence of change of conditions upon amount recalled, J. EXPER. PSYCHOL., 1931, **14**, p. 632.

ANTICIPATION AND ERRONEOUS RESPONSES *

BY F. H. LUMLEY

Ohio State University

I

In previous papers [1] errors made in serial learning were analysed and found to be anticipatory. The errors for whole groups of subjects were put together and the anticipatory relationships determined with reference to these combined errors. The purpose of this paper is to present additional evidence for the importance of the anticipatory type of error, and to show that it occurs in individual cases as well as in the massed data for a group of subjects.

Anticipation in maze learning is best defined in terms of the culs-de-sac which the subject enters. It means that the subject tends to go more frequently into those culs-de-sac which correspond with the direction or turning of the true path a short distance ahead.

The errors analysed in this paper were made by human subjects learning finger mazes. The mazes were learned under the supervision of Dr. Scott and Mr. Henninger of Ohio University in connection with another problem.[2] The maze patterns used were of the T type and the Miles X type. Several different patterns were used, and the number of

* Dr. T. C. Scott and Mr. L. L. Henninger of Ohio University graciously permitted the use of their data for analysis of the types of errors made in learning mazes. In the tables of this article their original system of maze coding has been followed and the results for the individual mazes may be compared with those given in their report. I am also indebted to Dr. Scott and Mr. Henninger for suggestions given during the preparation of this article.

[1] Lumley, F. H., An investigation of the responses made in learning a multiple choice maze, *Psychol. Monog.*, 1931, No. 189, **42**, 61 pp.; Anticipation of correct responses as a source of error in the learning of serial responses, J. EXPER. PSYCHOL., 1932, **15**, 195–205; Anticipation as a factor in serial and maze learning, J. EXPER. PSYCHOL., 1932, **15**, 331–342.

[2] Scott, T. C., and Henninger, L. L., The relation between length and difficulty in motor learning; a comparison with verbal learning. J. EXPER. PSYCHOL., 1933, **16**, 657–678.

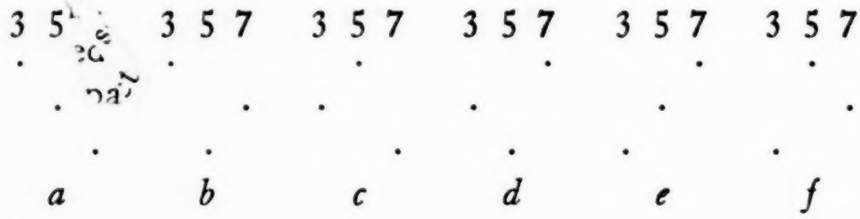
culs-de-sac varied from 8 to 15. Some of the subjects received special instructions, in addition to the ordinary preliminary information. The ordinary instructions resembled those usually given to maze learners. The subjects with additional instructions were shown a maze of the same unit structure as the one to be learned. They were told to learn the maze by memorizing moves and counting the successive moves in one direction, so that when they had learned a maze they would have a series of numbers in mind.

In accordance with the method of anticipation analysis previously explained, an error made by entering a cul-de-sac which has the same direction as the true path one turn ahead is called an error with one degree of anticipation. The entrance of a cul-de-sac that has the same direction as the true path two turns ahead is called an error with two degrees of anticipation, and so on. It is easy to see that a particular cul-de-sac might have the same direction as the true path at several different places further on in the maze. In such cases the cul-de-sac is always credited with anticipation for the nearest part of the true path that has the same direction. Other additional anticipations are disregarded. Defining anticipations in terms of one, two, three, etc., degrees is purely arbitrary, as is also the method of determining the degrees of anticipation shown by culs-de-sac for any particular maze. Thus the term 'anticipation of three degrees' is only a statement of procedure for the particular maze in question and does not have general significance.

Most of the mazes in this study contained culs-de-sac which permitted only errors of three, five, or seven degrees of anticipation. The average number of mistakes made by a subject on those culs-de-sac showing three degrees of anticipation was computed, similarly the average number of mistakes on those culs-de-sac showing five degrees of anticipation was determined, likewise for seven degrees. Thus a single subject would have three different averages for errors made on culs-de-sac, namely for those culs-de-sac showing three, five, and

seven degrees of anticipation. In order to simplify the reading of the results these three averages were reduced to a single trend. Previously it had been shown that *the greater the degree of anticipation, the fewer the number of errors.* This is, therefore, one relationship which could exist between the three averages. To fulfill the conditions of this relationship the average for three degrees of anticipation would have to be greater than that for five degrees of anticipation, and the average for five degrees would have to be greater than that for seven degrees.

If all cases of equal averages are excluded it is mathematically possible to have six different relationships among the three averages. These relationships are shown schematically below. The numbers at the top refer to the degree of anticipation.



The letters given below each group are simply of use for easy designation of the group. The case italicized above, where anticipation varies directly with the degree, is shown by (*a*). The reverse of the above relationship would be case (*e*). We can regard (*a*) and (*b*) as the only possibilities which fulfill the conditions of the italicized relationship to any extent. (*a*) does this more perfectly than (*b*). Since the relationship between number of anticipation errors and the degree of anticipation has been found to be logarithmic rather than straight, it follows that the difference between three and five is more important for the rule than the difference between five and seven. Because (*b*) shows three greater than both five and seven, it is considered that the reversal of five and seven in (*b*) is of minor importance. We may say, therefore, that (*a*) and (*b*) follow the rule, (*c*) and (*d*) are neutral, and (*e*) and (*f*) are contrary to the rule or reversed. These terms will be used in the tables. To determine the importance of

anticipation in these maze experiments it was only necessary to classify each subject under one of the six different relationships. To exclude equal averages a half credit was given to each of the nearest relationships for subjects where these equal averages occurred. In some of the mazes the position of the culs-de-sac made only two different degrees of anticipation possible. Subjects who learned these mazes were classified in only two groups, according to whether their errors followed the rule or were reversed.

TABLE I

RELATION OF MAZE ERRORS TO ANTICIPATION

The Number of Subjects in Each of Six Groups, Classified According to the Type of Errors Made in Maze Learning

Maze				Relation of Errors to Anticipation					
Type	Additional Instructions	Culs-de-sac	Code	Follow Rule		Neutral		Reversed	
				a	*b*	*c*	*d*	*e*	*f*
T	Yes	10	WC	8	6			½	½
T	No	10	WC	16½	7½				1
T	No	10	WC	20	5				
T	Yes	15	X	4	2		1		
T	Yes	15	Y	6	1		1		
X	Yes	10	X-Y	1½	4½	2		½	½
X	No	10	X-Y	5½	6	2	½		
X	No	10	Y-X	6	3	2	3	1	
Total for above mazes				67½	35	6	5½	2	2

Table I shows the results obtained for several different mazes. It is read as follows (first row): On a 10 culs-de-sac T type maze subjects were given additional instructions. The errors for eight of these subjects showed an (*a*) type of sequence. Six of the subjects made errors showing a (*b*) type of sequence. One subject made errors showing a mixed (*e*) and (*f*) type of sequence. Therefore, fourteen subjects in learning this maze made errors which followed the anticipation rule more or less closely, while only one subject made errors of a type directly opposed to the rule. From inspection of the results for the individual mazes, as well as from the total figures, it is evident that the errors for most of the subjects

followed the anticipation rule. The rule is—the greater the degree of anticipation, the fewer the number of errors.

TABLE II

RELATION OF MAZE ERRORS TO ANTICIPATION

The Number of Subjects in Each of Two Groups Classified According to Errors Made in Maze Learning

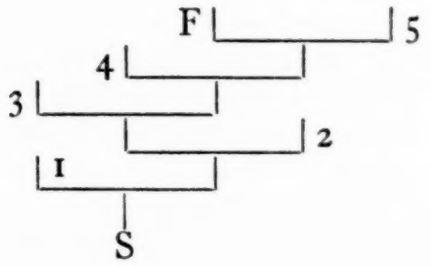
Maze				Relation of Errors to Anticipation	
Type	Additional Instructions	Culs-de-sac	Code	Follow Rule	Reversed
T	Yes	12	WC	14	1
T	No	12	WC	15	0
X	No	8	3–3	13	1

Table II supplements Table I and shows the results for three additional maze experiments where only three and five degree anticipations were possible. This table is read in the same manner as Table I, and shows that the errors for most of the subjects followed the anticipation rule.

Naturally anticipation is not a causal explanation of the subjects' errors. It is a descriptive explanation. It simply points out the fact that subjects make errors which the observer can classify in a certain way. It seems to the observer as if the subjects are doing something now which they should do in the future; in other words they are anticipating. They turn in a given direction before it is time to turn in that direction.

Other descriptions have been given of the kind of errors which subjects make. One such description states that human subjects tend to alternate in their choice of maze alleys.[3] In alternating they take a left turn, then a right turn, and then a left turn, etc. It is therefore interesting to take culs-de-sac where the degree of anticipation is the same, and determine the number of errors made on culs-de-sac where alternation leads to error and where lack of alternation leads to error. This may be made clear from the following diagram.

[3] Husband, R. W., A comparison of human adults and white rats in maze learning, *J. Comp. Psychol.*, 1929, **9**, 361–377.



Cul-de-sac 2 anticipates the turning of the true path by three degrees according to an arbitrary scheme used in previous reports and followed in this. The subject who enters cul-de-sac 2 does so by making the same turn by which he avoided entering cul-de-sac 1 and remained on the true path. Therefore, he is not alternating his right and left choices but choosing right each time, or making the same choice. We can call this the 'same type of turn.' Cul-de-sac 4 also anticipates the turning of the true path by three degrees. In avoiding cul-de-sac 3 the subject turned right. In order to enter cul-de-sac 4 he must turn left at the next place of choice. If he does so he is alternating his right and left choices, or making the opposite choice. All the culs-de-sac in the maze can be classified according to whether the subject makes the same or opposite turns in entering them.

In Table III the subjects have been classified according to the number of errors they made on the two types of culs-de-sac. Only errors on those culs-de-sac showing three degrees of anticipation were classified on the basis of alternation and non-alternation. This was done in order to maintain the same degree of anticipation in the comparison. Table III is read as follows (first row): Subjects were given additional instructions in learning a 12 cul-de-sac T type maze. Twelve subjects made more errors when the cul-de-sac turned in the opposite direction from the previous turning of the true path. These subjects tended to alternate. Three subjects made more errors when the cul-de-sac turned in the same direction as the previous turning of the true path. These subjects tended to keep going in the same direction, in other words, not to alternate. The results show that for the T type of maze subjects tended to alternate; for the X type of

TABLE III

ANTICIPATION ERRORS ON CULS-DE-SAC WHICH WERE, WITH REFERENCE TO PREVIOUS TURNS OF TRUE PATH (*e.g.*, LEFT), (1) THE SAME TYPE OF TURN (LEFT), (2) THE OPPOSITE TYPE OF TURN (RIGHT)

The Number of Subjects in Different Groups, Classified According to the Type of Error Made

Maze				Type of Turn	Number Subjects Errors Greater	Relation of Errors to Anticipation					
						Follow Rule		Neutral		Reversed	
Type	Additional Instructions	Culs-de-sac	Code			*a*	*b*	*c*	*d*	*e*	*f*
T	Yes	12	WC	Opposite	12	14				1	
				same	3	14				1	
T	No	12	WC	Opposite	15	15				0	
				same	0	12				3	
T	Yes	15	X	Opposite	6	4	2		1		
				same	1	2	3	1			1
T	Yes	15	Y	Opposite	8	6½	1	½			
				same	0	1½		4½		1	1
X	No	10	X–Y	Opposite	6½	7	5	½	1½		
				same	7½	6	5½	1½	1		
X	No	10	Y–X	Opposite	5½	7½	3	½	3½	½	
				same	9½	5½	6½	2½	1	½	

maze there was no pronounced tendency. The spacial characteristics of the T and X type of maze are different. In the T type the subject is often led away from the median line between the start and the finish. In the X type of maze, the subject remains very close to the median line. Therefore, the alternation shown with the T type of maze may be an expression of a tendency to return to the median line.

If alternation tendencies are found in the T type maze learning, it might be possible that these would account for the strength of the three degree anticipation errors. The relation between the various degrees of anticipation can be computed using only alternating or non-alternating culs-de-sac. Table III presents such an analysis. In the first row it is shown that when only those culs-de-sac (three degrees anticipation) having the opposite turning from the previous turning of the right path are taken, fourteen subjects showed

errors following the anticipation rule. One showed a reversal. When only those culs-de-sac having the same turning as the previous turning of the right path are taken the result is the same. Fourteen subjects showed errors following the anticipation rule, and one showed a reversal. From the other figures in the table it is evident that the anticipation rule holds whether the errors on the alternating or non-alternating culs-de-sac are taken in computing the third degree anticipation average. However, there were, as expected, not quite as many subjects with errors fitting the anticipation rule when only the non-alternating culs-de-sac are considered.

The figures given above show that anticipation in individual cases appears to be independent of alternation habits. Additional evidence for this conclusion may be obtained from an analysis of Husband's [3] data for rats and humans. The results obtained by Husband are particularly important because, as previously mentioned, he found that human beings tend to alternate whereas rats maintain a direction. The maze used by Husband (Warden U type T choice) had ten culs-de-sac. For purposes of the present analysis errors on the first and last culs-de-sac could not be used. The first cul-de-sac can show no alternation, since no choice precedes it. The last cul-de-sac can show no anticipation, since the true path stops at this point. The remaining culs-de-sac were divided into those showing three, five, and seven degrees of anticipation. The culs-de-sac were also divided into those showing alternation and those showing maintained direction (non-alternation). The following table (Table IV)

TABLE IV

THE AVERAGE NUMBER OF ERRORS MADE BY HUMANS AND RATS ON DIFFERENT TYPES OF CULS-DE-SAC

(*From Husband's Data*)

	Alternating Degree of Anticipation		Non-alternating Degree of Anticipation		
	3	5	3	5	7
Code numbers of culs-de-sac	7, 11, 17	5	13	9, 15	3
Humans	6.6	5.3	4.3	1.8	1.3
Rats	6.1	4.6	8.6	5.6	7.9

gives the average number of errors made by rats and humans on the different types of culs-de-sac, as well as the code numbers of the culs-de-sac (using Husband's notation) falling under each type.

Table IV is read as follows: Culs-de-sac 7, 11, and 17 were of the alternating type with three degrees of anticipation, cul-de-sac 5 of the alternating type with five degrees of anticipation, etc. Humans made on the average 6.6 errors on the alternating type of cul-de-sac with three degrees anticipation and 5.3 errors on the alternating type of cul-de-sac with five degrees anticipation.

This table shows that alternation is a contributing factor to the amount of error, but that anticipation results are evident in spite of the presence of alternation. Apparently the two factors work cumulatively and do not neutralize each other in this maze.

In saying that the culs-de-sac are alternating or non-alternating, or in saying that they show three degrees of anticipation, nothing is attributed to the cul-de-sac as such. The terms are only used to signify the relation of the subject to the culs-de-sac. The terms are a convenient method of reference.

Factors such as primacy and recency and the order of learning exert an influence on the results obtained with reference to anticipation. They may neutralize the action of anticipation or may reinforce it. In both cases it would be important to determine the action of these other factors, since errors which gave results in accordance with the anticipation generalization might be only a result of primacy in learning. The influence of primacy was shown in most of the results for the different mazes. Table V gives the results for errors made on the first three culs-de-sac, where primacy might be expected to exert the strongest influence. By obtaining the total number of errors on each cul-de-sac for the group of subjects it can be determined arbitrarily by inspection whether the effect of primacy is strong or weak. The figures show that primacy influences the results and neutralizes anticipation when it is strong. Therefore, the anticipa-

TABLE V

RELATION OF ANTICIPATION TO ERRORS MADE ON FIRST THREE CULS-DE-SAC

The Number of Subjects in Different Groups, Classified According to Errors Made in Maze Learning

Maze				Effect of Primacy	Relation of Errors to Anticipation					
					Follow Rule		Neutral		Reversed	
Type	Additional Instructions	Culs-de-sac	Code		*a*	*b*	*c*	*d*	*e*	*f*
T	Yes	12	WC	Weak		12			3	
T	No	12	WC	Medium		12½			2½	
T	Yes	10	WC	Strong	1½	2	2	1	1½	5
T	No	10	WC	Medium	10½	4	12	2	4½	8
T	Yes	15	X	Medium	3	2				
T	Yes	15	Y	Medium		1½	2	½	½	1½

tion results given in the preceding tables are more significant, since primacy is not a contributing factor to their appearance.

Table V is read as follows (row 3): Instructed subjects on a 10 culs-de-sac T type maze were distributed in all six groups on the basis of errors made on the first 3 culs-de-sac of the maze. The errors made by two subjects placed them in group (*b*) and the errors made by five subjects placed them in group (*f*). The effect of primacy in the case of the subjects learning this maze was strong.

In this connection, Hull's analysis of Lepley's conditioned response learning theory is of much interest. Hull deduces that the effects described by the terms *primacy-recency* and *anticipation* are merely two aspects of the functioning of conditioning (formation of conditioned responses) in serial learning.[4]

II

It is important to discover whether we can find examples of anticipation in the everyday experiences of people, and in types of experiment which are not primarily psychological. Recently a radio survey was made of the city of Omaha. Over fifteen hundred persons were called by telephone and asked questions concerning the program to which they were listening at the time of the call. The exact time of making

[4] Hull, C. L. The delayed conditioned reflex and serial learning, *Psychological Review*, 1934.

the interview was noted, so it was a simple matter to determine whether the program mentioned by the person was really on the air at the time in question. In a few instances the programs mentioned by the listeners had either occurred earlier in the evening or had not yet been broadcast. Naturally the number of cases available for analysis was not very great. Many people did not have their radios turned on. A considerable proportion of listeners could not name the program which they were hearing. And finally, the replies of the listeners were often correct.

In all, some 61 erroneous program mentions were received; 26 of these anticipated a program to be given later; 35 referred to a program already broadcast. The relation between the number of misplaced programs mentioned and the number of minutes between the actual scheduled time and the time of the interview is given as follows:

Type of Reference	Length of Time in Minutes between Interview and Program							
	01–14	15–29	30–44	45–59	60–74	75–89	90–104	105–up
Number of Anticipations	6	6	8	0	3	0	3	1
Number of Backward References	13	8	8	5	1	0	0	0

This table shows that 8 listeners said they were hearing programs which they could not have heard until 30 to 44 minutes after the time of the interview; 5 listeners said they were hearing programs which had already been given 45 to 59 minutes previously. It will be seen that backward reference as well as anticipation is shown in the listener's radio program information.

III

An opportunity to determine the influence of anticipation in 'natural' circumstances occurred in connection with a broadcast radio talk. Dr. W. R. McConnell broadcasts regularly two weekly geography lessons for the Ohio School of the Air over WLW in Cincinnati. Last year the Bureau of Educational Research in coöperation with the Payne Fund conducted a series of tests on these broadcast talks. Teachers

in different parts of Ohio, Kentucky, Indiana, and West Virginia were furnished with test blanks to give before and after the broadcast talks. The pretest was different from the end test.

One of the talks, which was about seventeen minutes in length, contained the following passage:

Now how far is Australia from the rest of the world? It is nearly 10,000 miles from New York City to Australia by way of the Panama Canal. Trace that route on your map. It is 12,000 miles from England to Australia by way of the Suez Canal and 14,000 miles from England to Australia by way of the Panama Canal. From Australia northeast to San Francisco the distance is 6,500 miles, north to Japan 3,000 miles, west to Africa 5,000 miles, east to South America 7,000 miles. Australia is nearer Asia than any other continent, but say, it's 2,500 miles from India. Australia is as far from Asia as England is from North America.

It is evident that if someone were to learn the above passage in serial form the associations would be as follows:

Distance	Australia	New York City	10,000
"	"	England via Suez Canal	12,000
"	"	England via Panama Canal	14,000
"	"	San Francisco	6,500
"	"	Japan	3,000
"	"	Africa	5,000
"	"	South America	7,000
"	"	India (Asia)	2,500

In this sense the problem corresponds to a paired associates learning problem. The first term is given and the second is to be reinstated. From such experiments we should expect that the first term would tend to reinstate terms which occur later in the series as well as the right terms. Now in the geography broadcast situation the material was presented only once. The pupils had no idea that it could be considered a series of numbers. They were intent only upon making associations between the first and second terms. In fact, since the pre-test for pupils whose end tests were used in this analysis contained no questions on numbers, the pupils may have ignored the numbers in the talk. The test itself was given in multiple choice and completion form. Only the results for the completion test will be reported here. This particular test was given to 117 pupils as a pre-test and to 97 others as an end-test. The completion test gave *no* suggested answers.

Combining 12,000 with New York City may be considered an anticipation of one degree. An anticipation of two degrees would be shown by saying that the distance between Australia and San Francisco is 5,000 miles. A backward reference of one degree would be evident if the pupils replied that Australia was 3,000 miles from Africa.

TABLE VI

AVERAGE NUMBER OF ERRORS PER PUPIL PER QUESTION SHOWING VARIOUS RELATIONSHIPS

(IN THOUSANDTHS)

Degrees of Relationship	Test and Relationship							
	Pre-test		End-test		Pre-test		End-test	
	Back Reference	Anticipation	Back Reference	Anticipation	Order of Size		Order of Size	
					Lower	Higher	Lower	Higher
6	21	26	30	0	21	0	8	0
5	10	16	20	17	24	3	8	8
4	5	15	38	6	18	7	13	30
3	12	16	31	27	27	8	21	41
2	15	17	23	28	7	8	21	37
1	12	9	28	26	22	11	22	39
0	11		185		11		185	

Table VI gives the answers in terms of anticipation and backward reference after an appropriate correction has been made for the number of possible anticipations and backward references. Table VI is read as follows (first row): Column one. Each pre-test pupil made on the average 21 thousandths errors by selecting a number which occurred six places back in the series with reference to the right number. Column two. Each pre-test pupil made on the average 26 thousandths errors by selecting a number which occurred six places further ahead in the series with reference to the right number. Columns three and four read the same way for the end-test. Column five shows that 21 thousandths errors were made on the pre-test by selecting a number six degrees less in size than the right number. (For example, choosing 2,500 when 12,000

is the right answer.) Column six shows that no errors were made on the pre-test by selecting a number six degrees greater in size than the right number. Column seven shows that 8 thousandths errors were made on the end-test by selecting a number six degrees less in size than the right number.

The other rows are read in similar fashion. Naturally where there is zero anticipation, zero backward reference, and zero order of size, the figures show the number of correct answers given by the pupils. On the end-test this amounted to 185 thousandths per pupil per question. In other words 18 percent of these numerical questions were answered correctly.

Table VI shows that on the pre-test no sign of the anticipation generalization was evident. Neither was there any relationship between the degree of backward reference and the amount of error. In other words, pupils selected answers almost completely by chance and there was no evidence of association between the numbers. On the end-test however the expected relationship between degree of anticipation and the number of errors appeared. Although the relationship was not very regular, it is clear that as we pass from one degree of anticipation to six degrees the number of errors becomes much less. No relationship between the degree of backward reference and the amount of error was shown on the end-test.

In order to determine the relationship between size of the erroneous number chosen and size of the correct number the figures given in columns five through eight of Table V were computed. These figures show that on the whole pupils more often chose numbers which were near the correct number in size than far away from it. If the correct number happened to be 6,500 the pupil was more apt to choose 7,000 than 10,000 as the proper answer. In the case of these figures the degree of relationship is taken from the list of numbers given in the text on Australia. For example, 5,000 is considered to be one degree below 6,500; 12,000 is three degrees above 6,500. These 'order of size' figures seem to show that the pupil classifies the numbers at least vaguely according to their size, although he may not be able to remember the

exact distance between Australia and New York. They also show that possibly the anticipation results given in the same table may be due to size discriminations, since size also enters into the serial order of the numbers in the text. This is most noticeable in the case of the first three numbers, namely 10,000, 12,000, and 14,000. However, when these first three numbers are disregarded the data for the remaining five numbers give results which are practically the same. Anticipation is still shown. Presumably relationships between the size of the numbers are not responsible for all of the anticipation relationship noted.

It is interesting to observe that a difference in distance estimation exists in comparing the pre- and end-test results. On the pre-test the pupils tended to underestimate the distance. On the end-test they were more apt to overestimate the distance and give mileages which were too large. This measures in a way the impression which Dr. McConnell created in his talk. He tried to establish the idea that Australia was a far-away and lonely continent. From the pupil responses there is indication that he was quite successful in implanting this idea.

In the radio talk on Australia other numbers were given besides the ones contained in the quoted paragraph. Because the numbers in the quoted paragraph were given near together and because they were all of the same type (*i.e.* miles) a separate analysis was made of them. In the whole seventeen minute radio talk, however, 34 numbers were mentioned. These numbers referred to miles, years, people, degrees, inches, feet, etc.: (*a*) 9 of the numbers modified miles, (*b*) 8 modified years, (*c*) 8 modified people, (*d*) and the rest modified degrees, inches, etc. The errors or wrong numbers given by pupils in answer to completion questions were analyzed with reference to the above four classes of numbers. This was done in two ways.

First, the numbers were only considered with reference to the particular group in which they fell, *e.g.* numbers modifying people. Anticipation was shown when another number modifying people and occurring further on in the talk was sub-

stituted for the right number. Suppose the talk contained these numbers: 10,000 people, 2,500 inches, 5,000 people, 40 degrees, etc. If 5,000 was substituted for 10,000 an anticipation of one degree would be shown for numbers in the same group. However, if 2,500 was substituted for 10,000 this would be an intergroup substitution. No attention was paid to such substitutions in computing the results according to the first of the two ways. No substitution outside of the same group of numbers was considered as showing anticipation or backward reference.

Second, the distinction between the groups mentioned above was disregarded. All substitutions of numbers irrespective of the words they modified were considered. In the above case putting 2,500 in place of 10,000 would be an anticipation of one degree, putting 5,000 in place of 10,000 an anticipation of two degrees, and 40, in place of 10,000 an anticipation of three degrees.

TABLE VII

AVERAGE NUMBER OF ERRORS PER PUPIL PER QUESTION SHOWING ANTICIPATION RELATIONSHIPS (IN THOUSANDTHS)

Degrees of Relationship	Test and Relationship							
	Numbers of the Same Kind				All Kinds of Numbers			
	Anticipation		Backward Reference		Anticipation		Backward Reference	
	Pre-T	End-T	Pre-T	End-T	Pre-T	End-T	Pre-T	End-T
6	5	0	7	8	8	2	10	10
5	3	6	5	7	4	4	5	7
4	4	9	3	17	4	5	11	27
3	6	9	4	13	9	6	5	10
2	5	9	6	17	10	12	4	13
1	4	15	6	18	3	13	6	16

Table VII shows the results when all 32 numbers in the talk are considered and the errors made in reporting them are analyzed. 'Numbers of the same kind' refers to the first way of computing anticipation. 'All kinds of numbers' refers to the second way. From this table it is clear that anticipation is shown on the end-test regardless of the way in which the

substitutions are treated. No sign of the anticipation generalization is apparent from the pre-test results. All figures in this table have been corrected for the number of possible anticipations for the different degrees.

Backward reference is only slightly shown on the end-test. Theoretically if any extraneous factor (such as number size) caused the appearance of anticipation it should also cause the appearance of backward reference. This, therefore, is a slight indication that the anticipation results are genuine.

IV

In summarizing this paper the following tentative conclusions may be stated: (1) Anticipation in maze learning is shown in the case of individual subjects as well as for massed data. (2) Anticipation in maze learning is shown whether the results are computed from errors made on alternating culs-de-sac or on non-alternating culs-de-sac. (3) Radio listeners anticipate the programs they are going to hear over the radio. (4) Pupils' memory for numbers given in a radio talk is in accordance with the anticipation generalization.

(Manuscript received December 5, 1932)

PUNISHMENT BY ELECTRIC SHOCK AS AFFECTING PERFORMANCE ON A RAISED FINGER MAZE

BY MILTON B. JENSEN

The subjects in this investigation were all college students, members of a class in experimental psychology. None of them had had previous experience with maze learning of this type though all were familiar with the pencil mazes of the Michigan Non-Verbal Series.

The mazes were constructed of No. 11 double pointed tacks driven into three-ply paneling. Runways along the true path are all the same length (the width of four tacks) and distances between tacks at the end of each section are all the same, thus insuring identification of the true path only by trial Errors are distinguishable by tacks standing alone as opposed to the series of four on the true path. Each choice on the maze has three alternatives: right, left and straight ahead. Correct choices are distributed so as to constitute no discernible series.[1] Three-choice rather than two-choice patterns decrease chance factors in the learning situation.

Fifty-five subjects learned Maze 'A' consisting of 30 choices. The maze was considered learned on the trial last preceding three consecutive errorless trials. Each subject was blindfolded and traced the path with the index finger on the preferred hand. The other fingers on that hand were bandaged to prevent their usage. Instructions were to learn the maze as rapidly and with as few errors as possible. Further talking by the experimenter was avoided as far as possible. The subject was permitted to verbalize at will.

[1] Maze 'A' runs in a major direction from left to right. The course along the true path, beginning at the starting point, is (R = right, L = left, S = straight ahead): L, R, L, R, S, L, L, S, R, R, S, L, R, L, R, R, S, L, L, R, R, S, R, S, L, R, L, S, L, S.

Maze 'B' runs in a major direction from right to left. The course along the true path is: S, L, R, L, S, R, L, L, S, R, L, R, S, S, L, L, S, R, S, R, R, L, L, R, S, L, L, R, R, L.

Backtracking was prevented by telling the subject when he started backwards. Learning was completed at one sitting and under typical laboratory conditions—no attempt was made to eliminate the customary noises. Forty-four of the subjects were paired as nearly as possible on the bases of: sex, total errors, trials to learn, errors on the first trial, and average rate of movement in the maze. Analysis of data for Maze 'A' is given later in this paper. The 22 pairs of subjects learned a second maze ('B') similar to 'A' and presumably of equal difficulty, half of them with punishment by electric shock each time an error was made, and their peers on the above enumerated items, with conditions identical except that no punishment was inflicted.

Correlations between performance on Mazes "A" and "B" for the 22 "control" subjects were: Errors, .92; Trials to learn, .89; Rate, .90. These data (though 22 constitutes a small population) together with the judgments of those acquainted with the study (particularly Professor C. Spearman and the author) indicate that the two mazes are approximately comparable learning situations. No data are available as to the relative difficulty of the two mazes since all the subjects learned Maze "A" first. Marked facilitation in the subsequent learning of Maze "B" will be noted from the data presented in Table I.

The electric shock was applied through the first two

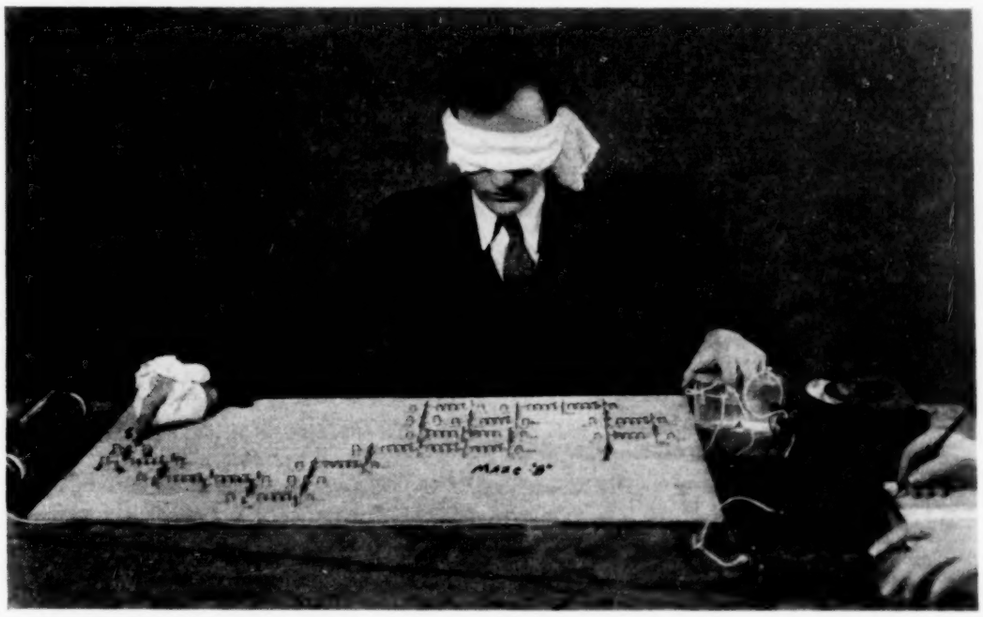
TABLE I

DATA FOR CONTROL AND EXPERIMENTAL GROUPS ON BOTH MAZES

(Group Averages)

	Total Errors	Trials to Learn	Rate	Errors on First Trial
Maze 'A':				
Control	208.14	22.90	.93	26.09
Experimental	209.59	22.68	.99	25.23
Percent difference	.70	.22	.70	3.30
d/σ_d	.04	.08	.82	.44
Maze 'B':				
Control	92.18	14.64	.78	20.05
Experimental	145.22	19.59	.94	20.77
Percent difference	57.50	33.80	21.00	3.60
d/σ_d	2.08	2.59	1.93	.44

fingers on the hand not used in tracing the pattern of the maze. These were immersed to the second joint in small jars of water containing terminals from a salt-bath rheostat connected to a regular lighting circuit. Application of electric current through liquid terminals is the only way of insuring constant voltage that I have found practicable. With this apparatus variations were not discernible on a meter calibrated in volts.



Subject learning Maze "B" under experimental conditions. *A*, salt bath rheostat; *B*, key switch for closing circuit when errors were made; *C*, voltmeter; *D*, resistance coil. Position of the experimentor is indicated at the right. Time for each trial was recorded by means of a stop watch.

Punishment was applied by the experimenter closing the circuit as soon as possible after the subject touched one of the tacks off the true path. It was in the neighborhood of .12 second in duration—the approximate reaction time of the experimenter.

The voltage used as punishment was determined for each subject in a manner analogous to that employed with the Whipple Pressure-Pain Balance. Each subject was acquainted with the apparatus. His fingers (the first two on the non-preferred hand) were immersed in jars of water with no current passing through them. The voltage was increased slowly and at a uniform rate (as far as the experi-

menter was able) until the subject said it hurt. This procedure was gone through from five to ten times on different days with each subject. The average of the voltages taken before complaint on the last three trials was applied as punishment. For the 22 subjects (11 men and 11 women) these were: 10, 11, 18, 19, 20, 20, 20, 21, 23, 25, 25, 27, 27, 27, 28, 30, 32, 32, 32, 42, 43, 50. Subjects evidencing fear (so far as I could determine) were not placed in the experimental group.

Conditions for controls and experimentals in learning Maze 'B' were apparently comparable with the exception of the punishment. The controls had their first two fingers on the unused hand immersed in water, though no current was applied.

Data relative to performance of the control and experimental groups on both mazes are given in Table I. Data from performance on Maze 'A' compare the groups on the initial learning. Data for Maze 'B' show, presumably, the effects of the punishment. Marked effects are seen in at least errors and trials.

In examining data of this type with respect to precision of measurement we may use the term reliability despite the glibness and indefiniteness generally associated with it. I have described accuracy in terms of correlations between halves of apparent comparability treated by the Spearman-Brown formula. The halves were rendered comparable by procedures I described in 1928.[2] These coefficients for 55 subjects on Maze 'A' are given in Table II.

TABLE II

RELIABILITY COEFFICIENTS FOR PERFORMANCE ON MAZE 'A'

	r_{11}
Errors (odd vs. even numbered items)	.938
Errors (odd vs. even numbered trials)	.990
Trials to learn (odd vs. even numbered items)	.980
Rate (means, odd vs. even numbered trials)	.958

Just how accuracy of measurement should be described is beyond the province of this paper.[3] The data presented

[2] *Genetic Psychology Monographs*, May 1928, p. 375.

[3] So far, efforts of psychologists at describing accuracy have failed to separate the consistency or accuracy of the instrument used from the variability of the performance

indicate accuracy, consistency, or reliability (whichever you prefer) of a nature satisfactory for investigations of this type.

Data from this study are in terms of errors, trials to learn, and rate of movement. These are, it seems to me, the major elements of maze performance as ordinarily measured, though I would not intimate that they are the only ones that can be measured. Casual observation indicates that time may not be used as a descriptive term discreet from errors since it is appreciably a resultant of the distance traversed—a factor appreciably dependent upon the number of errors made. If then, errors are to be used in describing performance in the maze, time may not be used. I have used mean rate of movement for each trial and for each subject. Correlations between these elements (errors, trials, and rate) for Maze 'A' (Table III) demonstrate discreetness of function.

TABLE III

CORRELATIONS BETWEEN ELEMENTS MEASURED IN PERFORMANCE ON MAZE 'A'
(Corrected for Attenuation)

	Total Errors	Trials to Learn	Mean Rate
Total Errors		.824 ± .029	−.030 ± .090
Trials to Learn	.824 ± .029		.296 ± .083
Mean Rate	−.030 ± .090	.296 ± .083	

Rate as used above is the average over the entire learning period. Quite different results are obtained from correlating errors and rate on the individual trials. This evidences that change in rate is a pertinent characteristic of the learning process ($r = -.596 \pm .014$).

More adequate methods of describing maze performance would contribute materially to psychological procedures. Analysis of variability of performance, it seems to me, has marked promise. Techniques involving moments around the mean are unsatisfactory, however, due to skewness in the distribution of data. Examination of the voluminous litera-

of the subjects. Until, and unless, experimentation provides the data necessary for this separation, statements as to accuracy of measurement in psychology must be made with this in mind.

ture in this field reveals considerable looseness of terminology and inadequacy of interpretation. Reliability coefficients obtained by correlating errors on odd and even numbered items is described as 'best' by one prominent investigator, presumably, because it gave a higher value than any other he determined—a naive procedure, to say the least.

While errors, trials, and rate are taken as basic elements in the present study, there are sufficient weaknesses in them as usually used to make questionable the findings reported from every investigation involving the maze as a so-called 'trial-and-error' method of learning. It will be recalled that errors made on the first run through the maze were used in pairing subjects for the present study. In terms of probabilities, the subjects, on the first run through the maze, should make errorless choices on approximately one-third of the items. The data show them as making errorless choices almost 50 percent of the time on the first run through both mazes. Acquaintance with the experimental procedure leads me to the conclusion that learning of this type is not strictly a chance situation at the outset. Table IV gives the means and the significance of differences between them for the first trials of the control and experimental groups on both mazes.

TABLE IV

DATA RELATIVE TO FIRST RUNS OF BOTH GROUPS THROUGH THE MAZES
(Each maze has 30, three-choice items—60 errors in all)

	Errorless Items	Items With One Error	Items With Two Errors	Total Errors
Maze 'A':				
Control mean	13.95	10.55	7.32	26.09
Experimental mean	14.09	12.05	6.59	25.23
d/σ_d	.11	1.39	.78	.44
Maze 'B':				
Control mean	14.87	11.04	4.09	20.05
Experimental mean	14.72	9.78	5.55	20.77
d/σ_d	.14	1.66	2.12	.44

Obviously, these results are not in accord with normal expectancy. The range of errors on the first trial through Maze 'A' was 34 (4 to 38) and through Maze 'B,' 24 (6 to 30). Even though errors on the first trial through Maze 'A' corre-

late closely with those on subsequent trials (.685 ± .048) we may not conclude that this obviates the necessity of comparability of chance factors in such learning situations. The relationship described as .685 between errors on initial and subsequent trials on Maze 'A' may mean that later performance is predicated upon chance factors affecting the initial trial, or that chance is not as great a factor on the first trial as has been assumed by most investigators. Whether one of these or some other explanation is found factual, comparability of chance factors is essential to experimentation in this field. It will be remembered (Table I) that this was under control so far as group comparisons are concerned. Unfortunately this was not so adequately equated in the pairing of subjects. Until, and unless, the effects of chance factors are, or can be determined, it is imperative that they be equated in comparisons. Unless their effects are known, learning experiments must eliminate them or hold them under control.

This study has shown marked influence of punishment by electric shock upon errors and trials and possibly upon rate of movement under the conditions described. It is suggested that, in studies involving the maze as a learning situation, only one element of performance be investigated at a time. With errors, time, and rate under consideration, this would mean the holding of two constant and observation of affective changes in the third. The influence of chance must be known or eliminated if the maze is to be a strictly scientific instrument. Furthermore, analysis must be made to determine what the various elements of maze performance actually predict. A suggestion in this direction is given in the following correlations (Table V) between total errors, trials to learn, and mean rate of movement in Maze 'A' with eight other measures: McCall Multi-Mental Test, Willoughby Emotional Maturity Scale, simple visual reaction time, grade point average, and score on the Features Discrimination, Aiming, Tapping and Mazes of the Michigan Non-Verbal Series by Greene. These show distinct relationship (negative) between trials and McCall Multi-Mental Test score ($-.91 \pm .02$) and

TABLE V

CORRELATIONS BETWEEN ERRORS, TRIALS AND RATE ON MAZE 'A' AND EIGHT OTHER VARIABLES

	Errors	Trials	Rate
Errors *		.82 ± .03	−.03 ± .09
Trials *	.82 ± .03		.30 ± .08
Rate *	−.03 ± .09	.30 ± .08	
McCall Multi-Mental	−.14 ± .10	−.91 ± .02	.14 ± .10
Willoughby EM Scale	−.29 ± .09	−.18 ± .10	−.01 ± .10
Simple reaction time (visual)	.03 ± .10	.06 ± .10	.00 ± .10
Grade point average	−.17 ± .10	−.84 ± .05	−.21 ± .09
Feature Discrimination (Greene)	−.15 ± .10	.09 ± .10	.32 ± .09
Aiming, right hand (Greene)	.07 ± .10	−.14 ± .10	.09 ± .10
Tapping, right hand (Greene)	−.02 ± .10	−.04 ± .10	.08 ± .10
Mazes (Greene)	.05 ± .10	.17 ± .10	.15 ± .10

* Corrected for attenuation.

grade point average ($-.84 \pm .05$). Low or zero relationships are seen between errors, trials and rate and the other variables. The correlations given in Table III are repeated here. They are the only ones which are corrected for attenuation.

I am deeply indebted for advice as to procedure and treatment of data to Professors Joseph Peterson and Charles Spearman. Without countless hours of gratuitous labor on the part of my former students at Western Kentucky State Teachers College, this study could not have been made.

(Manuscript received December 28, 1932)

THE EFFECT OF PUNISHMENT DURING LEARNING UPON RETENTION

BY LELAND W. CRAFTS AND RALPH W. GILBERT
New York University

The specific object of the present experiment was to determine the effect upon the retention of a maze of the administration of electric shock as punishment for error during the learning. Since, however, such punishment is presumably an incentive, a motivating condition, for human subjects, the aim of the experiment was equally to compare the retention of a maze learned under two different degrees of motivation: (1) that normal and usual for college students who volunteer for psychological experiments, and (2) the above, plus whatever additional motivating influence electric shock may possess.

Numerous experiments dealing with the effect of many kinds and degrees of motives upon learning in both animals and human beings have of course been performed. But very few investigators have concerned themselves with the influence of such factors upon retention. Dodson [1] found that light discriminations were not only better learned but also somewhat better retained after a 21 day interval by rats whose only incentive during the learning had been electric shock for incorrect choices than they were by groups who, being hungry in varying degrees, had been given food for correct choices. Valentine,[2] however, found no consistent differences in retention after an interval of 90 days between rats who learned a maze with and without shock, this in spite of the great superiority of the punishment groups in the learning. But these two experiments are obviously inadequate to permit any generalizations as to the effect of differences in motivation

[1] J. D. Dodson, Relative values of reward and punishment in habit formation, *Psychobiology*, 1917, 1, 231–276.

[2] R. Valentine, The effect of punishment for errors on the maze learning of rats, *Jour. Compar. Psychol.*, 1930, 10, 35–54.

during learning upon later retention. And furthermore, prior to the completion of the present experiment, no investigation had appeared in which there had been any study of this problem with human subjects.[3]

We have of course abundant data to show that the learning process is always facilitated by the presence of appropriate motivation (aside from certain disruptive effects attributable to too intense stimulation). Indeed it is quite possible that without such motivation only 'latent learning' occurs.[4] It is, however, quite certain that we cannot assume a priori that an act whose learning has been accelerated through the presence of some unusually effective motive would necessarily be better, or even as well, retained as would one that had been learned under less stimulating conditions. For the more motivated learning ordinarily involves less time and fewer trials than does the less motivated, and so is precisely that in which frequency of reaction, amount of practice, number of repetitions, is least. And while for learning the significance of frequency has become more and more doubtful—as Kuo[5] has effectively argued, it may well determine, not selection, but at most the degree of fixation of responses selected on some other basis, *e.g.* effect—its importance for retention might be quite otherwise. For the long retention of an already selected response might well be dependent on the degree of fixation of that response, and this fixation might in turn be closely correlated with its prior frequency. Certainly the studies of over-learning[6] have in the main shown that the more practiced is the better retained even though not always proportionately so.

Therefore the question as to the effect upon later retention

[3] The present experiment was completed before the publication of M. E. Bunch and F. D. McTeer's, The influence of punishment during learning upon retroactive inhibition, this JOURNAL, 1932, 15, 473–495.

[4] H. C. Blodgett, The effect of the introduction of reward upon the maze performance of rats, *Univ. Cal. Publs. Psychol.*, 1929, 4, No. 8, 113–134.

[5] Z. Y. Kuo, The nature of unsuccessful acts and their order of elimination in animal learning, *Jour. Compar. Psychol.*, 1922, 2, 1–27.

[6] For example, the two experiments of W. C. F. Krueger, The effect of overlearning on retention, this JOURNAL, 1929, 12, 71–78, and Further studies in overlearning, *ibid.*, 1930, 13, 152–163.

of varying degrees of motivation during the learning becomes a legitimate and, we believe, an important experimental problem. If the correct responses, though more quickly selected because of greater motivation, are less well retained, then such motivation during learning becomes detrimental to retention and something of advantage may thus be sacrificed when some unusual incentive produces rapid mastery of a problem. If, however, the more motivated learning is as well or better retained than the less motivated, then a decreased frequency during learning, if brought about by increased incentive, becomes as valuable for later retention—and so presumably for fixation—as a greater frequency associated with an incentive of less influence.

The present experiment is intended to contribute to the solution of this problem in so far as its specific conditions may be considered to permit.

METHOD

Subjects.—The subjects were 50 male and 50 female undergraduates at Washington Square College, New York University, assigned to two groups of 50 each, the experimental and the control, in such a way as to equate them for sex, class of origin, and approximately for college year and chronological age. None of the subjects had had any previous experience in maze learning.

Apparatus.—The maze employed was McGeoch and Melton's 'medium maze,'[7] cut from three-ply veneer, 8.5 by 11 in., and mounted on a bakelite base. Brass contact plates were inlaid at the ends of the ten culs-de-sac and each was connected with a common binding post by wires concealed beneath the maze. A Porter inductorium operating on three dry cells provided the shocking current. As it was necessary to exclude all auditory stimuli when the shock was given, the make-and-break vibrator was placed in another room. The shock was administered to the subject's active hand without the use of an extra electrode through a stylus having a bakelite handle in which were inlaid 16 brass rings 1/8 in. wide at intervals of 1/8 in. Alternate rings were connected across the secondary of the inductorium by flexible wires. The brass tip of the stylus was insulated from the rings on the handle and, being connected in the primary circuit of the coil, closed the latter whenever it was brought in contact with any one of the plates at the ends of the culs-de-sac. All stylus connections were made on the inside of the handle, and all wires braided to form a soft, flexible cord. The same stylus was

[7] J. A. McGeoch and A. W. Melton, The comparative retention values of maze habits and of nonsense syllables, this JOURNAL, 1929, 12, 392–414.

used by both experimental and control subjects, simple knife switches being provided to break the coil circuits.

Because of the irregularity of the oscillating current in the secondary circuit of an inductorium, the average voltage and amperage employed cannot be accurately stated. The former may however be estimated as ranging from 200 to 1000 volts, the latter as being only a small fraction of an ampere.

Curtains concealed both the maze and the experimenter from the subject.

Procedure.—The procedure was in broad outline as follows. The two groups, experimental and control, learned the maze to the criterion of two of three successive trials without 'contact' errors, *i.e.* without going so far into a cul-de-sac as actually to touch with the stylus the brass plate at its farther end. The experimental group received an electric shock whenever they made such an error, the control of course did not. After an interval of one week, the subjects relearned the maze to the same criterion as before. Since, however, the experiment was designed to measure the effect of punishment during *learning* upon retention, no shock was now employed with either group.

As to the procedure in detail: (1) Scoring was in trials, time and errors. Errors were classified into two types: (*a*) Contact errors, as defined above. Every separate contact with the end of a cul-de-sac was counted as one reror. Movement into a cul-de-sac which did not go so far as to cause contact with the end thereof was not scored as an error, although a record of such entraces was kept for the two final successful trials. (*b*) Retracing errors. Every change in direction, *i.e.* every turning of a corner, during a retracing, was counted as a separate error.

(2) The criterion of learning was, as we have said, to make two of three successive trials without contact errors. Hence a subject might make on these two trials any number of entrances and of retracing errors and still satisfy the criterion. The reason for this departure from the customary procedure in maze learning experiments was simply that since shock was administered for contact errors only, it seemed illegitimate to include in the criterion itself avoidance of any errors other than these. As the results will show, however, by the time a subject had eliminated all the contact errors, he had almost entirely abandoned errors of other types also. Any subject who could not learn the maze within 3600 sec. actual working time, or who showed little or no progress after about an hour in the laboratory, was classified as failed and his record discarded. The number of such failures was two males and six females in the experimental group and four males and two females in the control.

(3) Since the experimental subjects showed great individual differences in their sensitivity to electric shock, the amount to be used with any given subject was determined as follows: The subject repeatedly touched with the stylus a brass plate connected with the inductorium until—the amount of shock being gradually increased—he reported it to be 'too much.' It was then slightly decreased and the learning began with this amount, one which was therefore only a little less than the subject said he could stand and which was definitely annoying to him but not precisely painful. From a more objective viewpoint, the intensity used was sufficient to cause a visible negative response in the subject, *e.g.* a slight withdrawal or a stiffening of the fingers, but not enough to cause him to drop the stylus, nor to withdraw it so violently and spasmodically that later jumping out of the alleys of the maze could be foreseen. When in the course of the practice a subject gave evidence of decreasing sensitivity to the shock by a visible diminution in the speed and promptness of his withdrawal from the ends of the culs-de-sac, the amount was increased. Consequently almost every experimental subject completed his learning under the stimulus of a shock considerably greater than the one

with which he began. The intensity required for the male subjects was on the average greater than that employed for the female; some of the men seemed but little affected by the maximum amount which could be produced, whereas many of the women reacted violently to an intensity barely perceptible to the experimenter.

(4) A rest interval was allowed after each trial. Since, however, its length never exceeded about 60 sec., the exact duration depending on the length of the preceding trial, the practice may be termed massed.

(5) The subjects were requested to return a week after the learning, but were not informed that they would then be given a retention test. After the conclusion of the latter, however, they were asked whether they had expected such a test, whether they had made a drawing of the maze during the interval, and how much they had tried to rehearse or recall it. Their answers show that almost all of them had expected a retention test, that most of them had rehearsed the maze during the interval, and that a few had even attempted to draw from memory the true path. But since our records show that the amount of such rehearsal, etc. was approximately the same for the two groups, we believe it does not constitute a factor capable of exercising any differential influence upon them and that it may therefore be ignored.

(6) The subjects received a typewritten sheet containing the necessary explanations and instructions. These were, however, of the customary type and so need not be reproduced here.

RESULTS

The results for learning, as reference to Table 1 will show, are as follows: The experimental groups are consistently superior according to all of the criteria employed. The differences are, however, greater for trials and errors than they are for time and are much smaller with the male subjects. In the two final trials, in which the criterion of learning was satisfied by the absence of any contact errors, the experimental groups are slightly inferior in speed and in entrance errors, but the differences are throughout very small.

The results for retention are given in Table 2. In the recall score on the first trial of the relearning the experimental groups are superior in errors but inferior in speed and the differences are in all cases very small. In relearning the experimental groups are consistently superior according to all criteria (except that with the male subjects the average number of trials to relearn was 5.9 for the control group and 6.0 for the experimental), although the differences are statistically unreliable and are extremely small for the male subjects. It should be noted, however, that in no comparison save the one mentioned above does any control group surpass any experimental. In percent saved there are no consistent differences between the groups, although both the experi-

TABLE 1

LEARNING

Criterion	Experimental		Control		Diff./S.D. Diff.
	Aver.	S.D.	Aver.	S.D.	
Males					
Trials	16.2	8.83	19.9	6.05	1.71
Contact Errors	124.0	79.68	185.3	103.77	2.34
Retracings	133.3	134.50	166.5	119.46	0.92
Time	1268.0	645.24	1438.4	749.07	0.86
Two Final Trials,					
Entrance Errors	0.92		0.36		
Retracings	0.56		0.58		
Time	29.5		28.6		
Females					
Trials	18.2	7.89	26.7	9.67	3.40
Contact Errors	130.9	60.29	254.8	113.39	4.80
Retracings	138.0	81.02	271.0	138.00	4.15
Time	1608.2	580.60	2169.7	664.75	3.18
Two Final Trials,					
Entrance Errors	0.92		0.68		
Retracings	0.40		0.52		
Time	29.7		27.6		
Both Sexes Combined					
Trials	17.2	8.40	23.3	8.74	3.52
Contact Errors	127.4	70.74	220.0	114.10	4.89
Retracings	135.6	110.97	218.7	139.23	3.3
Time	1438.1	636.92	1804.0	797.00	2.54
Two Final Trials,					
Entrance Errors	0.92		0.52		
Retracings	0.48		0.55		
Time	29.6		28.1		

mental groups are slightly inferior in contact errors and slightly superior in retracings. In percent saved per trial, error and second of the original learning the experimental groups are in every case superior; the differences are especially large with the female subjects. In general, therefore, the various measures of retention employed either show the experimental groups to be superior, or fail to indicate any consistent difference between them and the control.

TABLE 2
RETENTION

Criterion	Recall		Relearning				Diff./ S.D. Diff.	Relearning, Percent Saved			
			Experimental		Control			Experimental		Control	
	Experimental Aver.	Control Aver.	Aver.	S.D.	Aver.	S.D.		Aver. Percent Saved	Percent Saved per Trial, etc. of the Learning *	Aver. Percent Saved	Percent Saved per Trial, etc. of the Learning
Males											
Trials			6.0	5.89	5.9	4.62	0.08	63.0	3.89	70.0	3.54
Contact Errors	4.9	5.6	16.7	16.27	16.8	14.57	0.02	86.5	0.70	90.9	0.48
Retracings	3.8	4.7	8.3	9.30	11.7	16.27	0.91	93.8	0.60	93.0	0.56
Time	59.1	54.1	198.0	160.95	214.3	166.44	0.35	84.4	0.067	85.1	0.059
Females											
Trials			5.3	7.02	9.4	7.78	1.95	70.9	3.90	64.5	2.42
Contact Errors	5.0	5.0	16.5	22.75	24.7	26.18	1.18	87.4	0.67	90.3	0.35
Retracings	2.8	3.8	6.4	6.82	19.0	31.25	1.97	95.4	0.69	93.0	0.34
Time	58.8	51.5	179.0	149.05	275.4	202.37	1.92	88.9	0.055	87.3	0.040
Both Sexes Combined											
Trials			5.7	6.49	7.6	6.62	1.49	66.9	3.89	67.2	2.98
Contact Errors	4.9	5.3	16.6	19.78	20.7	21.55	1.00	86.9	0.68	90.6	0.41
Retracings	3.3	4.2	7.3	8.21	15.3	25.00	2.14	94.6	0.63	93.0	0.45
Time	58.9	52.8	188.5	155.40	244.8	187.77	1.63	86.6	0.061	86.2	0.049

* Obtained by dividing the percent saved of trials, errors and time by the number of trials, errors and seconds of time respectively during the learning.

Discussion

(1) *Learning.*—That the use of electric shock as punishment for errors did accelerate the process of maze learning with our subjects is, we believe, clearly shown by the data of Table 1. The differences in favor of the experimental groups are not in all cases reliable, but they are present for all groups and for all criteria. Furthermore our data are in full accord with theoretical expectation and with the results of all other experiments, both with animals and human beings, in which such a stimulus has been employed. The large differences between the female groups are similar to those found for maze learning by Bunch [8] and by Bunch and McTeer,[9] and although the differences between the male groups are small, they compare favorably with those reported by Rexroad [10] for multiple choice reactions, by Vaughan and Diserens [11] for mazes, and by McTeer [12] for series of finger responses.

This superiority of the experimental groups may in theory be attributed either to the guidance or to the incentive value of punishment for error, or to both. That such punishment has guidance value in the sense that it serves to mark and define an error clearly for the subject is, we suppose, indisputable. Rexroad [13] has already noted the 'instructive' function of electric shock so administered, and the unsolicited comments of various of our subjects to the effect that during the retention test they 'missed' the shock and could no longer easily tell whether they were making errors or not is a further indication of it. But that such punishment also constitutes an incentive, a motivating factor, is, we believe, equally certain. Even a casual survey of the literature shows that such a function is assigned to it without qualification in

[8] M. E. Bunch, The effect of electric shock as punishment for errors in human maze learning, *Jour. Compar. Psychol.*, 1928, **8**, 343–359.

[9] M. E. Bunch and F. D. McTeer, *op. cit.*

[10] C. N. Rexroad, Administering electric shock for inaccuracy in continuous multiple choice reactions, this Journal, 1926, **9**, 1–18.

[11] J. Vaughan and C. M. Diserens, The relative effects of various intensities of punishment on learning and efficiency, *Jour. Compar. Psychol.*, 1930, **10**, 55–66.

[12] W. McTeer, A study of certain features of punishment in serial learning, this Journal, 1931, **14**, 453–476.

[13] Rexroad, *op. cit.*

current psychological opinion.[14] It likewise unquestionably falls within the definition of a negative incentive as given, for example, by Leuba,[15] who terms it a situation that stimulates the individual to exertions which usually persist until it has been avoided. Furthermore, this incentive value is suggested even more clearly and consistently than the guidance value is by the behavior, verbal and otherwise, of our subjects. Finally, the much greater superiority of the experimental group with the female subjects seems to require some reference to the motivating influence of electric shock for its explanation. For although it is possible that women profit more by guidance during maze learning than men do, there is nowhere any evidence to that effect. Whereas it is readily understandable that they really were, as they appeared to be, more afraid of the shock than the men were and that the punishment was therefore a stronger incentive for them.[16]

(2) *Retention.*—As the results given in Table 2 have shown, whether one is to argue that significant differences exist between the experimental and the control groups in retention depends somewhat on the criterion of retention employed. But since the experimental groups were superior according to the criteria of relearning and percent saved per trial, error and second of the learning, and were not consistently inferior to the control in either recall or absolute percent saved we may, we believe, conclude that on the whole they showed the better retention. This conclusion also agrees in the main with the results of Bunch and McTeer [17] who recently found that when a maze was learned with electric shock for error it was better retained, according to the criterion of relearning at least, after a six weeks' interval, and was also, according to all

[14] For example, experiments dealing with the effect of punishment for error upon learning are classified in reviews of the literature under the heading, the relation of motivation to learning. See J. A. McGeoch, The acquisition of skill, *Psychol. Bull.*, 1931, **28**, 413–466, and W. T. Heron, Habit formation and the higher mental processes in the rat, *ibid.*, 1932, **29**, 461–492.

[15] C. J. Leuba, A preliminary analysis of the nature and effects of incentives, *Psychol. Rev.*, 1930, **37**, 429–440.

[16] Since Bunch and McTeer, however, found no such divergence between their male and female groups, it is possible that the above difference is due to the accidents of selection only.

[17] Bunch and McTeer, *op. cit.*

criteria save trials, less subject to retroactive inhibition from practice on another maze interpolated during that period.

In explanation of this difference in favor of the experimental group we might, as in the case of the learning, refer both to the guidance and the incentive value of the punishment. It is questionable, however, whether recourse to the guidance factor alone would be adequate since, as Carr [18] remarks in his recent summary of guidance experiments, the latter show on the whole 'but little relation between the effects of tuition and the retention of the act.' Also it would seem unlikely that this factor could explain the much greater superiority of the female experimental group in retention any more than it could account for their similarly greater superiority in learning. From the viewpoint of the punishment as an incentive, a motivating factor, our data suggest, again in agreement with Bunch and McTeer,[19] that the employment of an additional incentive during the learning did improve the later retention. At least, the superiority of the experimental groups therein seems otherwise unexplainable. Furthermore it is in this way only that the greater superiority of the female experimental group seems understandable since, as have already pointed out, there is good reason to believe that they were the more motivated of the two punishment groups.

As to why additional motivation during learning should improve retention it is at present impossible to say, especially since the precise manner in which motivation influences learning itself is still highly problematical. Presumably electric shock induces what Leuba [20] terms the 'incentive attitude' with its increased muscular tension and more rapid and vigorous movements, and this more active attitude may, as Bunch and McTeer suggest, result in 'a more definite integration of the acquired activity.' Or it may be, as the latter also remark, that the decrease in errors and other unnecessary activity which the shock typically produces brings about "a greater difference between the functional strength of the acts 'selected' and that of the acts 'eliminated.'" In any

[18] H. Carr, Teaching and learning, *Jour. Genet. Psychol.*, 1930, 37, 189–219.

[19] Bunch and McTeer, *op. cit.*

[20] Leuba, *op. cit.*

case, however, if it was the more motivated and not merely the more guided learning which was the better retained, then the effect of the additional incentive was to increase the fixation of the correct responses to a degree exceeding that which the greater frequency of reaction in the less motivated learning could accomplish. And the validity of this conclusion would not be affected by inability to explain in just what fashion, neurological or otherwise, this greater fixation was produced. Furthermore the above conclusion is equally sound even when the retention of the more motivated groups only equalled that of the less motivated. For in such cases, since a less amount of practice with increased incentive was as favorable for retention as a greater amount without such incentive, the contribution to subsequent retention of every trial, error, and second of the learning was greater under the former condition than it was under the latter.

Summary

The aim of the experiment was to determine the effect upon the retention of a maze of the administration of electric shock as punishment for errors during the learning; or, in more general terms, to determine the effect upon its retention of the employment of an additional incentive designed to increase motivation during the learning process. The subjects were 100 college students divided into two groups of 25 men and 25 women each. Both groups learned the McGeoch-Melton Medium Maze to the criterion of two of three successive trials without error. The experimental group received an electric shock whenever the stylus came in contact with the end of a cul-de-sac. After an interval of one week the maze was relearned to the same criterion but without shock for either group. The results were: In learning the experimental groups were superior according to the usual three criteria of trials, errors and time. The differences were small, and greater for the female subjects, but were consistent throughout. In retention there was no significant difference between the groups in recall or in percent saved. In relearning, however, the male experimental group was slightly and

the female markedly superior to the control. It is therefore concluded that the punishment for error was advantageous not only for the learning of the maze but also for its retention. These results are in complete agreement with those of Bunch and McTeer and suggest that an act learned under more motivated conditions will be better retained—and so, presumably, better fixated—than one learned under less motivated, and this in spite of the fact that the greater incentive makes possible the attainment of the learning criterion with a less amount of practice and so with a decreased frequency of response.

(Manuscript received November 28, 1932)

VARIATIONS OF PSYCHOLOGICAL MEASUREMENTS ASSOCIATED WITH INCREASED VITAMIN B COMPLEX FEEDING IN YOUNG CHILDREN

(A Preliminary Study)

BY EVA RUTH BALKEN AND SIEGFRIED MAURER

Otho S. A. Sprague Memorial Institute and the Departments of Pathology and Psychology, the University of Chicago

Maurer and Tsai (**5**) have studied the effect of deficiency of vitamin B complex during the nursling period upon the maze learning ability of adult white rats. One group of animals was depleted while suckling depleted mothers, and another group, suckled by normally fed mothers, served as a control. They found that this early partial depletion of vitamin B complex was detrimental, the normally fed rats being about twice as efficient in maze learning as the depleted animals.

Hoobler (**3**) is of the opinion that there is a partial vitamin B deficiency in the diet of nursing mothers in America, as well as in the diet of infants artificially fed on cow's milk, sugar and cereal mixtures. Maurer and Tsai (**5**) have established the implications of such diet deficiency for animals, but hitherto there has been practically no adequate experimental data on its implications for human beings. Our investigation represents a first attempt to determine the applicability of the findings of Maurer and Tsai to human beings.

In this preliminary study, stress was placed on psychological and physiological technique and methodology and on ascertaining the desirability of continuing the study with new groups and adequate controls, rather than upon securing conclusive data for our general thesis. Following the method of McCollum (**4**) we sought a group of children who, because of

the straightened financial circumstances of their parents, had been presumably fed an inadequate diet while nurslings and were being fed an inadequate diet during the developmental years. We found such a group of children in one of the Chicago Public Schools.[1] The school was in one of the poorest sections of the city, where the families, largely of foreign nationality, have a record of having been given relief by several relief agencies for many years, even prior to the present financial conditions. The school has been supplementing this relief by maintaining a kitchen for the past sixteen years, where lunches have been furnished free or at a nominal cost. Recently, this venture has been financed mainly through the use of Chicago Tribune funds.

We found the diet of the children to be inadequate in both quality and quantity. The food consumed by these families consisted chiefly of coffee, rolls, bread of highly milled cereals, meat and potatoes. McCollum (4) and Davidson and Gulland (1) state that such a diet is deficient in vitamin B complex. The children were being fed this vitamin B deficient diet during their develop-years, and according to the belief of Hoobler (3) probably had been fed a similarly deficient diet while infants.

We selected 46 children, 24 boys and 22 girls, to serve as an experimental group. All were receiving free lunches at the school and these lunches comprised practically their entire sustenance for the school day. They ranged in age, at the beginning of the investigation, from 5.5 to 9.0 years, with an average of 6.6 years. Five were in the upper level of kindergarten, 41 were in grade 1. Eight were negro children, 38 were of either Mexican or Slavic nationality.

At the beginning of the study, these children were given the following tests: Stanford revision of the Binet scale;

[1] An outline of our general method of approach was submitted to Dr. D. P. MacMillan, Director of the Department of Child Study, for approval of its psychological aspects, and to Miss Florence Swain, Director of the Department of Home Economics, for approval of its nutritional aspects. Then with the approval and consent of the Board of Education, City of Chicago, we sought permission to carry out our investigation in the school referred to. This was very readily obtained through the coöperation of its principal, Mr. J. L. McBride, who placed the entire facilities of the school at our disposal.

Pintner-Cunningham Primary Mental Test, Knox Cube Test, Seguin Form Board, Whipple-Healy Tapping Test, Paterson's Five Figure Form Board, Young's Slot Maze A, Drawing of a Man, Casuist Form Board, Porteus Maze Tests, and Kohs Block Designs.

All of the tests were given by the one examiner, and individually, except in the case of the Pintner-Cunningham Primary Mental Test, which was given in groups of 5 to 8.

The testing was distributed over several sessions and given in the following order: Session I—Stanford revision of the Binet Scale; Session II—Knox Cube Test, Seguin Form Board, Whipple-Healy Tapping Test, Paterson's Five Figure Form Board, Young's Slot Maze A, Drawing of a Man, Casuist Form Board, Porteus Maze Test, Kohs Block Designs and, again, Knox Cube Test; Session III—Pintner-Cunningham Primary Mental Test. The testing in each session covered a period of about 30 minutes.

During the first four weeks of the experiment no additional vitamin B complex was fed. Beginning with the fifth week and continuing through the twentieth week the vitamin B complex was fed 5 days each week. The simplest procedure for its feeding without involving additional cookery was found to be in the addition of a vitamin B complex concentrate in the form of Yeast Foam Tablets and of Yeast Foam Powder [2] mixed with peanut butter to form a sandwich paste. The children were given each a minimum of 10 grams daily. The tablets and sandwiches were distributed by the examiner and consumed in her presence. Careful check was made upon the physiological reactivity of the children to the additional vitamin B complex. It was found that the children tolerated even more than the minimum allowance of 10 grams and without deleterious effects. It was only because of time spent in distributing the tablets and sandwiches and in seeing that it was eaten that the daily allowance was omitted.

While on the increased vitamin B complex concentrate diet the children were retested in the following manner:

[2] The Yeast Foam Tablets and Yeast Foam Powder used in this investigation were provided by the Northwestern Yeast Company, Chicago, Illinois.

Test	Number Retests	Total Number Tests	Time Interval Between Tests
Stanford revision *	2	3	10 weeks
Pintner-Cuningham	1	2	20 weeks
Knox Cube †	4	5	4 weeks
Seguin Form Board †	4	5	4 weeks
Tapping ‡	4	5	4 weeks
Five Figure Form Board †	4	5	4 weeks
Slot Maze A ‡	4	5	4 weeks
Drawing of a Man §	4	5	4 weeks
Casuist Form Board †	4	5	4 weeks
Porteus Maze †	4	5	4 weeks
Kohs Block Designs †	4	5	4 weeks

* For testing procedure and scoring see Terman, L. M., *The Measurement of Intelligence*, Houghton Mifflin Co., 1916.

† For testing procedure and scoring see Arthur, G., *A Point Scale of Performance Tests*, Commonwealth Fund Publications.

‡ For testing procedure and scoring see Bronner, Healy, Lowe, Shimberg, *A Manual of Individual Tests and Testing*, Little, Brown and Company, 1927.

§ For testing procedure and scoring see Goodenough, F. L., *The Measurement of Intelligence by Drawings*, World Book Co., 1926.

The results of this preliminary study with an experimental group of the effects of an increased vitamin B complex diet upon the performances in a series of mental tests of young children selected as being fed a diet inadequate in vitamin B complex are summarized in the following tables:

TABLE 1

STANFORD REVISION OF THE BINET SCALE

Series	Average Intelligence Quotient	Difference Between Averages, II–I, III–I	Average Gain
I	84.4		
II	87.1	2.7	
III	88.0	3.6	3.2

TABLE 2

PINTNER-CUNNINGHAM PRIMARY MENTAL TEST

Series	Average Intelligence Quotient	Average Gain
I	82.3	
II	84.0	1.7

TABLE 3
CASUIST FORM BOARD

Series	Average Score in Points	Difference Between Averages, II–I, III–I, IV–I, V–I	Difference Between Averages, II–I, III–II, IV–III, V–IV	Average Gain
I	.78			
II	1.46	.68	.68	
III	1.73	.95	.27	
IV	2.68	1.90	.95	
V	3.04	2.26	.36	.145

TABLE 4
KOHS BLOCK DESIGNS

Series	Average Score in Points	Difference Between Averages, II–I, III–I, IV–I, V–I	Difference Between Averages, II–I, III–II, IV–III, V–IV	Average Gain
I	.63			
II	1.20	.57	.57	
III	2.13	1.50	.93	
IV	3.11	2.48	.98	
V	3.54	2.91	.43	1.87

TABLE 5
KNOX CUBE TEST

Series	Average Score in Points	Difference Between Averages, II–I, III–I, IV–I, V–I	Difference Between Averages, II–I, III–II, IV–III, V–IV	Average Gain
I	2.94			
II	3.71	.77	.77	
III	4.69	1.75	.98	
IV	5.86	2.92	1.17	
V	6.54	3.60	.68	2.26

TABLE 6
SEGUIN FORM BOARD

Series	Average Score in Seconds	Difference Between Averages, II–I, III–I, IV–I, V–I	Difference Between Averages, II–I, III–II, IV–III, V–IV	Average Gain
I	26.13			
II	22.91	3.22	3.22	
III	22.36	3.77	.55	
IV	21.69	4.44	.67	
V	21.48	4.65	.21	4.02

TABLE 7
WHIPPLE-HEALY TAPP[illegible] TEST

Series	Average Score	Difference Between Averages, II–I, III–I, IV–I, V–I	Difference Between Averages, II–I, III–II, IV–III, V–IV	Average Gain
I	27.4			
II	30.8	3.4	3.4	
III	34.2	6.8	3.4	
IV	36.5	9.1	2.3	
V	39.1	11.7	2.6	7.8

TABLE 8
YOUNG'S SLOT MAZE A

Series	Average Score in Seconds	Difference Between Averages, II–I, III–I, IV–I, V–I	Difference Between Averages, II–I, III–II, IV–III, V–IV	Average Gain
I	120.7			
II	68.9	51.8	51.8	
III	50.6	70.1	18.3	
IV	40.2	80.5	10.4	
V	24.7	96.0	15.5	74.6

TABLE 9
PATERSON'S FIVE FIGURE FORM BOARD

Series	Average Score in Seconds	Difference Between Averages, II–I, III–I, IV–I, V–I	Difference Between Averages, II–I, III–II, IV–III, V–IV	Average Gain
I	257.6			
II	228.6	29.0	29.0	
III	195.5	62.1	33.1	
IV	157.9	99.7	37.6	
V	115.5	142.1	42.4	83.2

TABLE 10
PORTEUS MAZE TESTS

Series	Average Score in Years	Difference Between Averages, II–I, III–I, IV–I, V–I	Difference Between Averages, II–I, III–II, IV–III, V–IV	Average Gain
I	4.62			
II	5.14	.52	.52	
III	5.74	1.12	.60	
IV	6.80	2.18	1.06	
V	7.41	2.79	.61	1.65

TABLE 11

DRAWING OF A MAN

Series	Average Score in Points	Difference Between Averages, I–I, III–I, V–I, V–I	Difference Between Averages, II–I, III–II, IV–III, V–IV	Average Gain
I.....	6.3			
II.....	7.5	1.2	1.2	
III.....	7.7	1.4	.2	
IV.....	8.1	1.8	.4	
V.....	8.6	2.3	.5	1.7

Inspection of the foregoing tables shows improvement of performance in each of the battery of mental tests with the successive retestings, with each retesting as compared with the initial series and also with the immediately preceding series. The trend of the gains is so consistent that we are justified in concluding that the efficiency of performance by the young children, comprising our experimental group, in the given mental tests was increased following the initial series and while the children were being fed additional vitamin B complex. The data obtained from our present intensive study of a small group of children, do not permit us, however, to attribute the increased efficiency of performance entirely and directly to the increased vitamin B complex feeding. The improvement in performance may have been due to maturation, practice, or an improved neurological state arising from the feeding of additional vitamin B complex.[3]

We wish to stress again that the primary purpose of the present investigation was to develop physiologic and psychologic methodology and to determine the desirability and possibilities of an extended and controlled study of the general

[3] Tentatively, it appears that the improvement in the performance of at least one of these tests (Whipple-Healy Tapping Test) is not attributable to practice effects, as Gates (2) found that the control and practice groups obtained practically equivalent scores at the end of 18 weeks, during which time the control group had been tested during only the first and eighteenth weeks, while the practice groups were given daily practice in tapping throughout the eighteen weeks.

The gains in intelligence quotient obtained with the successive retestings (see Tables 1 and 2) may not be considered as significant inasmuch as several studies summarized by Terman (6) have shown that gains to the extent obtained in our study are to be expected upon retesting.

problem of the relation of vitamin B complex deficiency to mental measurements. The consistent gains noted in the foregoing tables indicate the necessity for an extended study of our general problem in order to determine which of the three possible facts already referred to was operative. Whether these gains are etiologically associated with neurological changes arising from the feeding of additional vitamin B complex will constitute the specific problem of another investigation now under way.

(Manuscript received January 9, 1933)

References

1. Davidson and Gulland, *Pernicious Anæmia*, The V. C. Mosby Co., 1930, p. 250.
2. Gates, A. I. and Taylor, G. A., An experimental study of the nature of improvement resulting from practice in a motor function, *J. Educ. Psychol.*, 1926, 17, 226–236.
3. Hoobler, B. R., Symptomology of vitamin B deficiency in infants, *Jour. Amer. Med. Assoc.*, 1928, **XCL**, 307–310.
4. McCollum, E. V. and Simmonds, N., *The New Knowledge of Nutrition*, New York, 4th ed., 507, 516, 112, and 517.
5. Maurer, S. and Tsai, L. H., The effect of partial depletion of vitamin B complex upon learning ability in rats, *Jour. of Nutrition*, Nov. 1931, Vol. 4, 507–516.
6. Terman, L. M., *The Intelligence of School Children.*

THE INFLUENCE OF COFFEE ON THE ASSOCIATION CONSTANT

BY GERTRUDE REIMAN

The Catholic University of America

I

The experiment about to be described developed from an attempt to analyze association by its equational constants (**3**).

The preliminary problem was the determination of a measure of associative learning based on an individual's complete performance and independent of any particular stage of learning. After a trial of various tests one was finally devised which was intended to limit the method of learning to a single type of procedure. The subjects were shown a series of ten cards on each of which was printed a letter and a two-digit number. Immediately after the cards had been presented the subjects were required to fill in the appropriate letters on a sheet of one hundred blanks which contained only the numbers. There were fifteen such sheets and the cards were exposed before each one, allowing fifteen repetitions for the learning of the ten letter-number combinations. From this test two series of values were obtained: the number of combinations associated and the number of blanks correctly filled. The first was believed to be a measure of associative ability and the second of what was termed facilitation, *i.e.*, the power to make ready use of the association formed.

The next step was to average the scores of the members of a fifth grade class to which the test was given, on the assumption that the curve of the average values might indicate the general type of curve to which the individual scores might correspond. The empirical curve seemed to show that the theoretical one should approach an asymptote, but it required further experimentation to determine which of the various forms of this type of curve should be selected. This was done by following Running's (**5**) method based on the differ-

ences of individual observations. In Moore's data it was found that as x (the successive learning periods) progressed arithmetically, the y differences became progressively smaller (the y values being the average number of associations established by the entire class in the xth trial).

Running gives laws for two equations, $y = a + bc^x$ and $\log y = a + bc^x$, which were both applicable in this case, but it was found by calculation that the latter of the two gave the closest agreement between observed and theoretical values and it was accordingly adopted.

It may be well to point out here that when the equation $\log y = a + bc^x$ is used to measure learning by the formation of associative bonds, the variables and constants have the following meanings:

y is a variable: the number of syllables learned in any one (the xth) repetition; it may also be the average number learned for all subjects or the average number for all trials of the same subject.

x is a variable: the number of repetitions, first, second, etc. to the nth repetition.

a is a constant which defines the task, that is, the number of items to be fixed by associative learning.

b is a constant which is always negative in learning experiments. It is a function of the inertia of the nervous system to be overcome by the process of learning. If the task is within the span of the subject so that everything is learned in one repetition then b is zero. The larger b the more the task exceeds the span of the learner.

c is a constant which defines the subject's ability to form associative bonds. Inspection of the equation shows that the smaller c is, the quicker the curve ascends towards its asymptote; that is, the fewer are the repetitions required for perfect learning.

Moore evolved a method of obtaining the constants of the equation $\log y = a + bc^x$ as follows: when a is known (*i.e.*, when the subject is given a definite number of associations to

be formed), if c is less than unity and $x = \infty$, then $\log y = a$; in other words, a is the logarithm of the total number of items to be associated. The following equations may be written from the observed values:

$$\begin{aligned} -bc &= a - \log y_1 \\ -bc^2 &= a - \log y_2 \\ -bc^3 &= a - \log y_3 \\ &\cdots\cdots \\ -bc^n &= a - \log y_n. \end{aligned}$$

From these equations the following ratios are given for the determination of c.

$$c = \frac{a - \log y_n}{a - \log y_{n-1}}$$

$$c^2 = \frac{a - \log y_n}{a - \log y_{n-2}}$$

$$c^p = \frac{a - \log y_n}{a - \log y_{n-p}}.$$

The average of the above values is obtained and we have thus a simple method of calculating c when a is known. The value of b may be obtained from the relation

$$\frac{\log y_n - \log a}{c^n} = b.$$

If one wishes to use the equation to determine the limit of 'facilitation' in the application of learned material then a is unknown. As already mentioned, the number of squares filled in is a measure of facile use of learned material. As a matter of fact we have used the maximum number of squares filled in as a measure of facilitation and designated this value as the speed-factor. It would be possible, however, to calculate a by taking as the y values the number of squares filled in after each repetition.

When a is unknown the calculation of c is almost as easy as when a is known. Here we have

$$\begin{array}{rl} \log y_2 &= a + bc^2 \\ \log y_1 &= a + bc \\ \hline \log y_2 - \log y_1 &= b(c^2 - c) \end{array} \qquad \begin{array}{rl} \log y_3 &= a + bc^3 \\ \log y_2 &= a + bc^2 \\ \hline \log y_3 - \log y_2 &= b(c^3 - c^2). \end{array}$$

From this we form the relations

$$\frac{\log y_n - \log y_{n-1}}{\log y_{n-1} - \log y_{n-2}} = \frac{b(c^n - c^{n-1})}{b(c^{n-1} - c^{n-2})} = c.$$

So also

$$\frac{\log y_n - \log y_{n-1}}{\log y_{n-2} - \log y_{n-3}} = c^2, \text{ etc.}$$

This relation holds as long as the difference between the subscripts in the numerator is the same as that between the subscripts in the denominator.

Since the subjects were children whose performance naturally fluctuated, the procedure of smoothing the scores by averaging them in threes was used and the extent of accidental variations thereby reduced; this, however, gives the data spuriously high reliability which must be borne in mind in interpreting probable errors.

In order to determine b when a is unknown the following relationship is employed:

$$\log y_{n+1} = \log a + bc^{n+1}$$
$$\log y_n = \log a + bc^n.$$

Hence

$$\frac{\log y_{n+1} - \log y_n}{c^{n+1} - c^n} = b$$

$$\frac{\log y_{n+2} - \log y_n}{c^{n+2} - c^n} = b, \text{ etc.}$$

The value for a is finally obtained by substituting the values for b and c in the observation equations.

The method outlined above is fortunately capable of simplification and improved accuracy. The average of the c values derived from all the powers of c that may be determined from the experiments is not the most perfect determination of c. Those obtained from the observation equations involving single differences of the y values are the most subject to accidental error; those from the double differences average by two's the accidental errors, etc. It was found by calculation that the average of c from the first to the nth difference becomes less variable until one comes to

the median value and then increases in variability to the *n*th value. Furthermore when the median *c* value is used the residual errors are less than when the average of all values is used. *B* also is best calculated from median values.

The proof of the equation $\log y = a + bc^x$ shows that the constant *b* involves *c* and another constant *d*.[1] Analysis of the equation indicated that the antilogarithm of *d* is a function of the inertia in learning, and *c* of the individual's power of overcoming this inertia; *a* is the logarithm of the number of items in the task or when the equation is used to measure speed it is the logarithm of the maximal attainable speed.

Experiments showed the value of *c* to be only relatively constant, but the validity of the equation is not thereby destroyed since it gives very close approximations in many and varied conditions.

It was thought that the new method of measuring association might lead to more clear-cut results in the study of the influence of various drugs on mental operations. The selection of coffee for this purpose was made for practical reasons: it is a drug-containing beverage in common use and represents the form in which that drug is generally taken. The psychological influence of caffeine or coffee has been studied by various investigators to the general conclusion that its effect is one of stimulation. The advantage of the present experiment, however, lies in the fact that the equational constants depend on the individual's entire performance from beginning to end and hence are better measures of ability than the result of learning at any single stage of its development. It was hoped that by determining the equational constants a less ambiguous result might be obtained than in some of the previous work dealing with the effect of drugs on human mental performance.

Earlier experiments (**4**) indicated that Moore's association test in its original form was too simple for advanced subjects so it was rendered more difficult by using nonsense syllables and three-digit numbers instead of single letters and two digits. The syllables chosen were taken from the lists given by Glaze (**2**). Ten forms of the test were prepared, the

[1] $d = b(1 - c)$. Cf. *Aspects of the New Scholastic Philosophy*, p. 219.

syllables and numbers being carefully selected and arranged to obviate any possible cues to forming the associations. Only those syllables were used which were pronounceable and all the syllables in a single form had the same associative value (as found by Glaze). It was thought that the increased difficulty of the syllable combinations over the letters might make it advisable to reduce their number so a preliminary investigation was made, using six to ten combinations, and it was found that eight was the most satisfactory number.

The syllables and numbers were printed on cards as for the simpler form and the booklets were made up of ten pages with one hundred spaces to the page. In this case, however, the syllables were given and the subjects were required to fill in the corresponding numbers as seen on the cards. The syllables appearing in the booklet were printed in chance order which differed for each page. In giving the test the syllables and numbers were shown in succession, each card being exposed only long enough for the experimenter to pronounce the combination, and the time allowed for each page was fixed, again after preliminary trials, at 150 seconds. The cards were shown in a different order for each page of the booklet.

The first attempt to determine the influence of coffee had to be discarded since it was found that ten forms of the test were too many and the drug effect was clouded by the practice effect. A second attempt was made with seventeen subjects, fifteen of whom were graduate students. This time only five forms of the test were used, A, B, C, D and E;[2] form A was given merely for practice and was not included in the results, while the succeeding four were given alternately without and with coffee. Three subjects found it impossible to finish the series, so that in their cases the data from only two tests were available.

The experiments were conducted on five different days at approximately the same hour, 4:00 P.M. On each of the four

[2] Forms A, B, C, D contained syllables of 33 percent associative value and form E those of 26 percent associative value, which may therefore be assumed more difficult to remember. If this were true, it would have slightly counteracted the effect of the coffee which was given before this test.

significant days the subjects refrained from drinking either tea or coffee before coming to the laboratory where, on the second and fourth days, they were given one cup of black coffee (prepared under controlled conditions) half an hour in advance of the test.

In scoring two counts were made: the number of combinations learned on each page and the number of spaces correctly filled in. The first record was used in determining the constants of association; the second denotes the speed-factor or facilitation. The individual's raw scores were smoothed by averaging them in threes and the constants c and b calculated in the manner already indicated, using median values.

The calculated scores (y_1, $y_2 \cdots y_n$) are secured by subtracting from log a the products of b multiplied by c, $c^2 \cdots c^n$. Y_0, which is a function of the individual's equipment prior to the first repetition in the learning series, equals log $a - b$; d is the product of b and $(1 - c)$, and also equals $y_1 - y_0$ (theoretical values).

In determining the results of the experiments the antilogarithms of b, d and y_0 were used as well as their logarithmic values.

Results.—When the values involved in the equation $\log y = a + bc^x$ had been calculated for each test the data for the no-coffee days were correlated with those for the coffee days and the sigmas of the differences found. The data included the speed-factor and y_1, the number of combinations learned on the first presentation.

TABLE I

	Mean$_1$ no coffee	Mean$_2$ coffee
speed-factor	43.548	57.935
y_1	1.6	1.97
log y_0	0.165	0.112
antilog y_0	1.649	1.574
c	0.706	0.586
log b	0.738	0.791
antilog b	6.367	9.835
log d	0.213	0.363
antilog d	1.711	3.670

Table I gives the means for the no-coffee and coffee days. It will be noticed that the largest differences are those shown

by the speed-factor, antilog b and antilog d; when, however, the formula

$$\sigma_{diff.} = \sqrt{\sigma_{av.1}^2 + \sigma_{av.2}^2 - 2r_{12}\sigma_{av.1}\sigma_{av.2}}$$

is applied and the ratio of the difference to the sigma of the difference is obtained it will be seen (Table II) that the elements showing the most significant differences are the speed-factor, c and log d.

TABLE II

	Difference Between Means	$\sigma_{diff.}$	$\frac{Diff.}{\sigma_{diff.}}$	Probability *	r
speed-factor	14.387	2.440	5.896	.9999999982	.675
y_1	0.370	0.203	1.823	.9658468	.269
log y_0	0.053	0.065	0.815	.7924609	.167
antilog y_0	0.075	0.197	0.381	.6483866	.037
c	0.120	0.026	4.669	.9999984940	.700
log b	0.053	0.071	0.746	.7721636	.154
antilog b	3.468	3.196	1.085	.8610361	.095
log d	0.150	0.052	2.885	.9980427	.425
antilog d	1.960	1.182	1.658	.9513399	.219

* Values taken from Table II: Pearson's Tables for Statisticians and Biometricians. London, 1923.

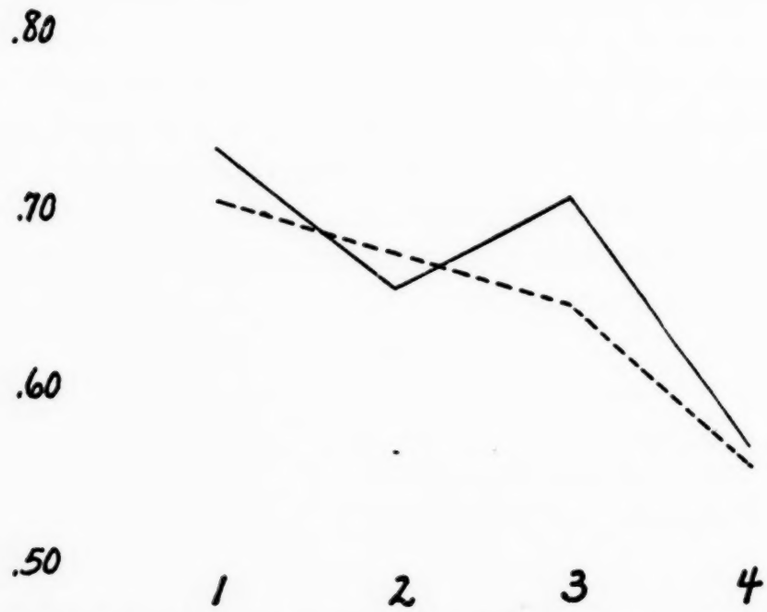


FIG. 1. The unbroken line shows the effect of coffee upon the average c values for each day; the broken line illustrates the practice effect.

Figure 1 illustrates the effect of coffee upon the constant c; it shows the fluctuation in the average c values of all the

subjects for the four days on which the tests were given. Remembering that the first and third days were without coffee and the second and fourth days with coffee, and that the lower the value of *c* the greater the facility of learning, we note that there is a marked decrease in the *c* value for the second and fourth day and a corresponding rise for the third day. The broken line, which illustrates the average *c* values in an experiment in which the tests were given on successive days to determine the practice effect (1), emphasizes the difference in learning ability with and without coffee.

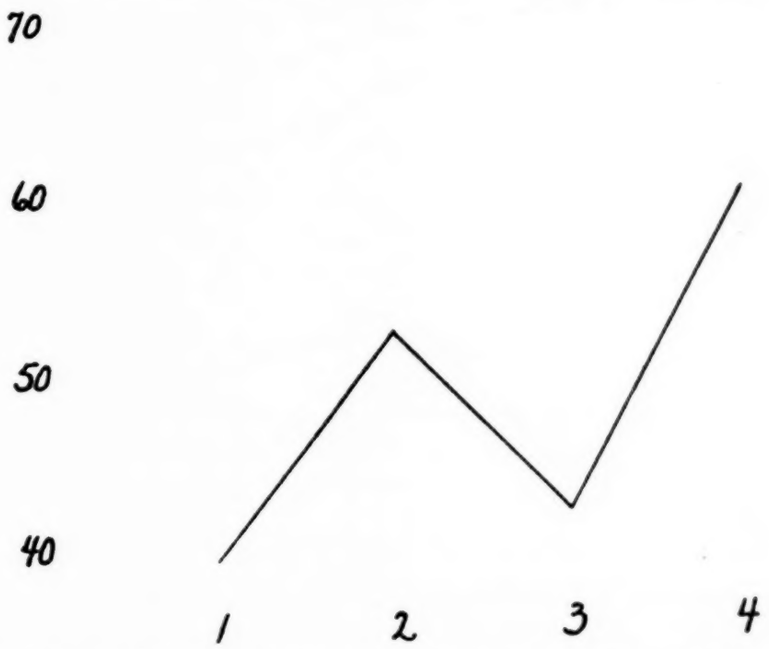


FIG. 2. The effect of coffee upon the average values of the speed-factor.

Figure 2 shows in like manner the change in the speed-factor and here the rise and fall corresponding to the absence and presence of coffee is even more pronounced. The subjects, with a single exception, succeeded in correctly filling more spaces after they had drunk coffee.

The average *d* values as shown in Fig. 3 exhibit a tendency to rise on coffee days, although the marked increase on the fourth day may be due in part to some unknown factor. It is interesting to note that the coffee stimulation which causes improvement in learning ability and speed tends also to raise, if it has any effect, the inertia of the nervous system.

The learning scores of all the individuals were averaged for the no-coffee and coffee days and the constants of the equation calculated for these mean values. The resulting curves are shown in Fig. 4 and illustrate still more clearly the improvement in learning due to the coffee stimulation. There is a close agreement between the theoretical and experimental results which is more noticeable for those without coffee. The theoretical curves were carried beyond the actual number of scores in the experiment in order to determine the relative

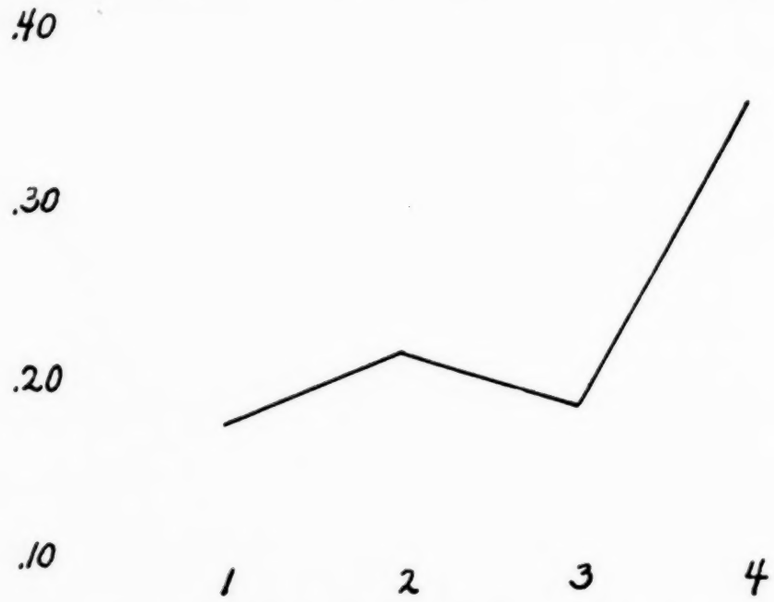


FIG. 3. The effect of coffee upon the average *d* values.

rapidity of their approach to the maximum, and here also the coffee values are decidedly superior.

Summary.—An attempt was made to investigate the influence of coffee upon the constants involved in a mathematical curve which may be used to express the conditions that obtain in associative learning.

The test used in this experiment consisted of eight combinations of nonsense syllables and three-digit numbers which the subjects were allowed ten opportunities to learn. Different forms of the test were given on four different days, alternately without and with coffee.

The equation which was applied to the data obtained from the test was: $\log y = a + bc^x$, where a, b and c are constants.

The constant a is the logarithm of the number of items to be associated and evidently defines the limit of the task.

The constant b represents in some manner the amount of work that must be done for all the items to be associated. It will be noticed that in learning experiments it is always negative and, when multiplied by c, represents an amount to be subtracted from a which is decreased with every repetition until the difference between a and bc^x becomes negligible. The decrease is brought about by multiplying b by a power

FIG. 4. The effect of coffee upon the average learning scores. The lower curve shows the results without coffee; the higher curve with coffee. Unbroken lines denote the empirical values; broken lines, the theoretical values.

of c. It should be noted that in learning experiments c is always less than unity. If b is very small to begin with the task is almost within the span of the subject. Therefore b is a function of the subject's span.

The constant b, however, equals $d/1 - c$. Theoretically, by giving arbitrary values to d, b can be made anything we please for any assumed value of c; practically, d is always a function of c. It was shown in Moore's study that $d = y_1 - y_0$ (the theoretical values). Although theoretically, given y_0, we may make y_1 anything we please, practically y_1 is always a function of c, so that neither b nor d is in practice independent of c.

The inertia of the nervous system relative to the task, *i.e.*, the block to perfect repetition which must be overcome in learning is

$$(\text{antilog})\ (\log a - \log y_0) = \frac{\text{antilog } a}{\text{antilog } y_0}.$$

For the normal or no-coffee series this inertia is 5.476; for the coffee series it is 6.181.

As indicated in Table II the difference between y_0 on the no-coffee and coffee days is not statistically significant, but if anything coffee tends to increase the inertia of the nervous system and therefore in this respect has a deleterious effect on learning. This double action may account for some of the conflicts in the literature. It should be remarked, however, that another piece of research (1) tended to show that practice, while lowering the c value, raises y_0 and therefore decreases the inertia of the nervous system. These facts would indicate that y_0 and c refer to different attributes of the nervous system.

It was found that coffee lowers the value of c and the amount of lowering is certainly significant. This lowering of c means an increase in the power of forming associative bonds.

Facilitation in writing the combinations learned was also improved by coffee, as shown by the increase in the value of the speed-factor on coffee days.

We may conclude, therefore, that this experiment shows without any ambiguity that the power to form associative bonds is increased by coffee.

References

1. Diggins, Julia. A Master's Dissertation: The Effect of Practice on the Constant of Association. The Catholic University of America, 1932.
2. Glaze, J. A., The association value of non-sense syllables, *Jour. of Genetic Psychol.*, 1928, **35**, 255–267.
3. Moore, T. V. The analysis of association by its equational constants, *Aspects of the New Scholastic Philosophy*, C. A. Hart, Ed. New York: Benziger Bros., 1932.
4. Reinhart, Sister Miriam. A scale for measuring the G-factors in intelligence, *Studies in Psychology and Psychiatry*. Baltimore: Williams and Wilkins, 1931, II, 5.
5. Running, Theodore R. *Empirical Formulas*. New York: John Wiley and Sons, Inc., 1917.

THE STATISTICAL LIMEN VERSUS THE AVERAGE AS A MEASURE OF VISUAL APPREHENSION

BY MILES A. TINKER

University of Minnesota

Examination of the investigations on range (span) of visual apprehension (often incorrectly called range of attention) reveals strikingly different and apparently conflicting results. Variation in experimental procedure and methods of evaluating data undoubtedly explain to a large degree these discrepancies. This makes it hazardous, of course, to compare the results obtained by the different experimenters. In general these data on visual apprehension have been presented in three ways: (1) as an arithmetical average with or without a measure of variability; (2) as a percent correct for each length of stimulus series used; and (3) as the statistical limen (with variability) which is a statistical refinement of (2). The statistical limen may be defined as the stimulus value for which there are correct judgments in 50 percent of the cases; *i.e.*, as frequently apprehended correctly as incorrectly. Fernberger and Dallenbach, who have completed or supervised considerable work in this field, favor the statistical limen for measuring range of apprehension (see Tinker, **4**).

The uncertainty involved in evaluating results of studies on visual apprehension should be mitigated to some degree when the effect of scoring methods on range is known. In an earlier investigation (**5**) averages derived by three methods of scoring were compared. The results showed that while absolute range (span) of apprehension was affected by method of scoring, range in relation to other ranges in the group was not materially altered. In the present study the statistical limen of visual apprehension will be compared with arithmetical averages derived in various ways.

The stimulus material consisted of a series of black letters $3\frac{1}{2}$ inches high, drawn with black ink on white cardboard, 7

by 28 inches in size. Only consonants were employed, and the series varied from 4 to 9 letters in length. Fifty stimulus cards for each length of series were constructed. When they had been used once the cards were shuffled and presented again so that 100 series of each length had been employed when the experiment was completed. There were two experimental periods per week for three weeks. One hundred stimulus series, an equal number of each length, were given at each session.

Seventeen university students, tested all in one group, served as subjects. The stimulus cards were exposed, one at a time, for a period of three seconds. Exposure was achieved by merely uncovering the stimulus card in a position clearly visible to all the subjects. Timing was done with a stop clock.

The subjects were instructed to view carefully the series of letters on the card while it was exposed. As soon as the stimuli were covered the students were to write down as many of the letters as they could remember and in the same order as on the exposure card. Omission of a letter was to be designated by a dash in the series.

Although it has been demonstrated that limens may be satisfactorily obtained from fewer observations, 100 stimulus series of each length were used in this study to insure stable averages. The statistical limen (Boring, 1) and averages for each length of stimulus series were computed for each subject. Group averages were also obtained. Probable errors were calculated for all averages and limens. The $P.E._L$ was computed with Culler's formula (2). The averages were computed in two ways: (1) from scores in which a credit of one was given for each letter reproduced in order, and in addition, a credit of one-half for each letter reproduced correctly but out of order; (2) from scores in which each letter correctly reproduced received a credit of one irrespective of whether it was or was not in order.

The computed means and limens, when the weighted score was given for letters out of order, are given in Table I. On comparing the limens in the last column with the means in

the preceding columns it is found that for every subject the limen is markedly higher than the largest mean. Apparently the absolute size of the statistical limen is not comparable to the average.

TABLE I

MEAN SCORE OF APPREHENSION FOR EACH LENGTH OF STIMULUS SERIES AND THE STATISTICAL LIMEN WHEN LETTERS OUT OF ORDER RECEIVED A WEIGHTED SCORE

Subject	Length of Stimulus Series						Limen
	4	5	6	7	8	9	
A	4.00	4.97	5.80	6.49	6.92	6.94	7.75
B	4.00	4.97	5.86	6.43	6.52	6.22	7.22
C	3.97	4.96	5.50	5.58	5.29	5.25	6.46
D	4.00	4.92	5.56	5.22	5.49	5.26	6.45
E	3.97	4.93	5.88	6.71	7.18	7.21	8.34
F	3.96	4.93	5.19	5.01	5.08	5.07	6.03
G	3.93	4.61	4.57	4.55	4.53	4.47	5.29
H	3.76	4.73	4.70	4.11	3.99	4.09	5.55
I	3.96	4.89	5.59	6.05	5.87	5.82	6.69
J	3.99	4.83	5.49	5.67	5.65	5.50	6.43
K	4.00	4.97	5.80	6.25	5.95	5.30	7.07
L	3.99	4.97	5.89	6.65	7.21	7.25	8.25
M	3.99	4.85	4.88	4.97	4.85	4.79	5.78
N	3.99	4.99	5.76	6.32	6.35	6.14	7.26
O	3.97	4.82	5.19	4.87	4.75	4.95	5.89
P	3.94	4.77	5.05	4.76	4.77	4.57	5.69
Q	3.98	4.87	5.23	5.46	5.41	5.46	6.31
Av. $P.E_M$	.01	.03	.05	.07	.08	.09	.04

There are other interesting trends in Table I. Lack of marked differentiation of means for stimulus series containing 4 or 5 letters indicates that these short series, which are apprehended almost 100 percent correctly, are of little or no value by themselves as measures of visual apprehension. Inspection of a subject's averages for series of 6, 7, 8 and 9 letters reveals three distinct tendencies in mean scores when the length of stimulus series increases markedly faster than the score: (1) the mean score increases slightly with each longer stimulus series (subject E), or increases through series 8 and then decreases at 9 (subject B); (2) the mean score holds fairly constant with increase in stimulus series (subject Q); and (3) the mean score decreases slightly with each increase in stimulus series (subject G). The responses of all subjects, therefore, were not affected in the same manner by a continued

increase in the length of stimulus series. Some subjects increased their scores, others were little affected, and still others were apparently somewhat confused by the longer series and their scores were shortened.

Comparison of the limens in the last column of Table I with the means in the four columns preceding it shows that the rank order of the limens is about the same as those of the means. The rank-order correlation between the limens and each of the columns of means yields coefficients ranging from .94 to .98 with an average of .965. Also the intercorrelations between the columns of averages range from .94 to .98 with an average of .963. These correlations justify the conclusion that if size of score in relation to others in the group is wanted, the limen and the averages are equally good measures of visual apprehension.

In the last row of the table are given the averages of $P.E._M$. The probable errors of the individual averages and limen revealed approximately the same trends as these averages. There was a gradual increase in the $P.E._M$ with increase in length of stimulus series. The $P.E._L$ was always about one-half that of the $P.E._M$ for series 8 and 9.

TABLE II

MEAN SCORE OF APPREHENSION FOR EACH LENGTH OF STIMULUS SERIES AND THE STATISTICAL LIMEN WHEN ALL LETTERS RECEIVED FULL CREDIT

Subject	Length of Stimulus Series						Statistical Limen
	4	5	6	7	8	9	
A	4.00	4.93	5.86	6.58	6.95	7.09	7.92
B	4.00	4.91	5.89	6.69	7.06	7.07	7.70
C	3.98	4.98	5.64	5.89	5.71	5.33	6.69
D	4.00	4.97	5.68	5.76	6.19	5.88	6.71
E	3.98	4.93	5.91	6.79	7.42	7.62	8.70
F	3.99	4.96	5.34	5.31	5.31	5.37	6.22
G	3.98	4.77	4.79	4.84	4.92	4.88	5.41
H	3.96	4.89	5.09	4.48	4.34	4.42	5.87
I	3.96	4.94	5.75	6.39	6.57	6.75	6.98
J	3.99	4.93	5.69	6.09	6.13	5.98	6.84
K	4.00	4.99	5.90	6.50	6.42	5.89	7.37
L	3.99	4.98	5.93	6.78	7.39	7.67	8.70
M	3.99	4.86	5.13	5.19	5.22	5.08	5.87
N	3.99	4.99	5.81	6.52	6.87	6.92	7.57
O	3.99	4.92	5.43	5.32	5.28	5.37	6.16
P	3.98	4.90	5.28	5.23	5.26	5.04	6.03
Q	3.98	4.90	5.46	5.62	5.72	5.62	6.38
Av. $P.E._M$	.01	.02	.04	.06	.08	.09	.05

The individual averages and the limens, when full credit was given for reproduced letters irrespective of order, are listed in Table II. Comparison of limens with the averages shows, as for the other method of scoring, that the limen is always decidedly larger than the largest average for that subject. In general the size of the individual averages tends to vary with size of stimulus series as it did with weighted scores. However, there is one important exception; there is a more pronounced tendency for a continuous increase in average with increase in length of stimulus series, with a rather marked increase for the longest stimulus series (9 letters). This indicates, of course, that with the longer series the tendency to reproduce letters correctly but in a wrong order increases.

As with the other method of scoring the correlations between the various averages (series 6 and above), and between the limen and averages are high: for the former the range is .94 to .98 with an average coefficient of .97; for the latter the range is .96 to .99 with an average of .98. Hence the limen and the averages are equally valid as measures of visual apprehension where score in relation to others in the group is desired.

In Table III an average apprehension score involving responses to all stimulus series (4 to 9 items) is compared to the statistical limen for each subject. The averages are, of course, smaller than several of the averages to the longer stimulus series (see Tables I and II), due to the small scores obtained on the series of 4 and 5 letters. Both averages and limens from weighted scores are slightly smaller than those from unweighted scores. The differences between averages and limens, however, are marked. When averages and limens are correlated with each other the coefficients range from .98 to .99 with a mean of .985. Again we find that an average score may be substituted for the statistical limen when one wishes merely to measure relative size of score.

Group averages for each length of stimulus series, for all series together, and for the limens are presented in Table IV. Where the stimulus series is somewhat longer than average

TABLE III

MEAN SCORES AND LIMENS OF VISUAL APPREHENSION. EACH MEAN AND LIMEN INVOLVES 600 RESPONSES TO STIMULI RANGING FROM 4 TO 9 LETTERS

Subject	Correct: Weighted Score		Correct: Any Order	
	Mean	Limen	Mean	Limen
A	5.86	7.75	5.91	7.92
B	5.66	7.22	5.95	7.70
C	5.09	6.46	5.29	6.69
D	5.08	6.45	5.42	6.71
E	5.98	8.34	6.11	8.70
F	4.89	6.03	5.05	6.22
G	4.44	5.29	4.69	5.41
H	4.23	5.55	4.52	5.87
I	5.37	6.69	5.73	6.98
J	5.19	6.43	5.47	6.84
K	5.38	7.07	5.62	7.37
L	6.00	8.25	6.13	8.70
M	4.72	5.78	4.91	5.87
N	5.59	7.26	5.85	7.57
O	4.76	5.89	5.05	6.16
P	4.64	5.69	4.95	6.03
Q	5.07	6.31	5.22	6.38

score (series 8 and 9) about one-half of an item more can be apprehended than in correct order. This agrees with Tinker's (**5**) findings. In both methods of scoring, averages increase consistently with increase in length of stimulus series through series 8. There is a drop in average score in series 9. The statistical limen of apprehension is, for each scoring method, approximately $1\frac{1}{2}$ letters larger than the averages of all (10,200) responses.

The above comparisons of the statistical limen with aver-

TABLE IV

MEAN SCORES FOR EACH LENGTH OF STIMULUS SERIES AND LIMENS OF VISUAL APPREHENSION FOR GROUP OF 17 SUBJECTS

Series	Correct: Weighted Score		Correct: Any Order	
	Mean	P.E.M	Mean	P.E.M
4	3.96	.003	3.99	.002
5	4.88	.007	4.93	.005
6	5.41	.013	5.56	.011
7	5.59	.018	5.88	.016
8	5.64	.020	6.04	.019
9	5.55	.021	6.01	.020
All	5.17	.006	5.30	.006
Limens	6.62	—	6.89	—

age scores of visual apprehension of letters show clearly that the absolute value of the statistical limen is always greater than the average irrespective of how the average is computed. The evidence adduced above also demonstrates that the statistical limen and any one of the various averages are equally good measures of apprehension for that subject when size of score in relation to other scores in the group is desired.

Fernberger (3) has stated that the statistical limen is a more reliable and more readily determined measure of range of attention (span of apprehension) than some others. For the kind of stimuli employed in this study (which are considerably different from those considered by Fernberger) it is extremely doubtful that the limen is any more reliable than the averages. Inspection of the above tables reveals very small variability of both averages and limens. Also Tinker (5) has shown that scores for apprehension of letters are very reliable, the coefficients ranging from .869 to .952. It is probable that the reliabilities of limens and averages as computed in this study are strictly comparable. Some of these individual averages are just as easy or easier to calculate than the limen. This is true especially when correctly reproduced items are given full credit irrespective of order.

The statistical limen as a measure of visual apprehension does have the marked advantage over the average in that it is unequivocally defined, and the definition undoubtedly has high validity since it is derived directly from the accepted psychophysical principle of the limen. It would appear that the choice of the statistical limen or the average as a measure of visual apprehension must be determined by the nature of the experiment and the use to which the data are to be put. If one is interested in span of apprehension from a theoretical viewpoint the statistical limen is obviously a more adequate measure. However, if span of apprehension is to be correlated with reading ability or the like, some form of average is probably to be desired because (1) an average score for each of a large number of subjects is easier to obtain than the limen, and (2) the perceptual situation from which an average is derived (*i.e.*, apprehending only part of a sentence at a glance)

is more comparable to the actual reading situation than an experiment for determination of the limen of apprehension. In certain experimental situations, such as measuring apprehension of series of dots, as in Fernberger's study (3), there is no choice, for the limen is the only adequate measure.

Summary and Conclusions

1. The statistical limen has been compared with averages computed in various ways, as a measure of visual apprehension of letters. Individual averages and limens for 17 subjects as well as group results were considered.

2. In every comparison the statistical limen was considerably larger than the averages. For the entire group of subjects, the limen was approximately 1½ items larger than the average.

3. The average scores of apprehension were highly correlated with each other (.963 and .970) and with the statistical limens (.965 to .980).

4. In general, scoring method produces a decided difference in absolute span of visual apprehension, but produces no practical difference in size of score in relation to other scores in the group.

5. The statistical limen as a measure of visual apprehension has an advantage in that it is unequivocally defined. However, the average appears to be a more adequate measure in certain studies involving correlation.

6. Choice of the statistical limen or the average as a measure of visual apprehension should be on the basis of suitability to the requirements of the experiment.

References

1. Boring, E. G., Urban's tables and the method of constant stimuli, *Amer. J. Psychol.*, 1917, **28**, 280–293.
2. Culler, E., Studies in psychometric theory: VI. The probable error of the limen and its derivation, *Psychol. Monog.*, 1926, **35**, No. 2, 85–90.
3. Fernberger, S. W., A preliminary study of range of visual apprehension, *Amer. J. Psychol.*, 1921, **32**, 121–133.
4. Tinker, M. A., Visual apprehension and perception in reading, *Psychol. Bull.*, 1929, **26**, 223–240.
5. Tinker, M. A., The effect of color on visual apprehension and perception, *Genet. Psychol. Monog.*, 1932, **11**, 61–136.

JUDGMENT TIMES FOR THE METHOD OF CONSTANT STIMULI

BY W. R. CARLSON, R. C. DRIVER AND M. G. PRESTON
University of Pennsylvania

The question of judgment times for the different categories of the psychophysical methods has recently been reopened by the work of Kellogg [1] and that of Fernberger and Irwin.[2] Kellogg confirmed the earlier findings of Martin and Müller [3] and of George,[4] all of whom found that the equality judgment required a longer time for its formulation than did the judgments of difference—either greater or less.

It would seem that these several studies are in complete conformity and that the question of the time relations of the different categories of the psychophysical judgment was rather definitely settled. But Fernberger and Irwin, working with the method of single stimuli, found no significant differences between the judgment times for the categories of 'less,' 'intermediate' and 'great.'

In view of this disparity of results for Fernberger and Irwin and all of the earlier studies, it has seemed worth while to repeat the experiment with the method of constant stimuli for several reasons. In the case of George's study, it would seem that there were too few reaction times taken for the averages to have real statistical significance. Let us hasten to add that the reaction times were in the nature of a by-product in this experiment inasmuch as George's primary interest was in obtaining introspective reports of the process of judgment.

[1] W. N. Kellogg, The time of judgment in psychometric measures, *Amer. J. Psychol.*, 1931, **43**, 65–86.

[2] S. W. Fernberger and F. W. Irwin, Time relations for the different categories of judgment in the 'absolute method' in psychophysics, *Amer. J. Psychol.*, 1932, **44**, 505–525.

[3] L. J. Martin and G. E. Müller, *Zur Analyse der Unterschiedsempfindlichkeit*, 1899, 1–96.

[4] S. S. George, Attitude in relation to the psychophysical judgment, *Amer. J. Psychol.*, 1917, **28**, 1–37.

George employed 5 *Os.* and obtained a total of only 2,529 reaction times. He also employed 13 categories of judgment so that the largest number of cases for all 5 *Os.* for any category of judgment is less than 700 reaction times. Kellogg's study, or at least that part which applies to the present paper, is based on only 1680 reaction times for his total of 5 *Os.* and for three categories of judgment.

Also the matter of the recording of reaction times is open to question in all of these studies. Martin and Müller used the very crude method of counting beats of a metronome, while George and Kellogg both employed graphic methods. Kellogg's time line,for example, was made by a Jacquet chronograph which marked intervals of only fifth seconds (200 sigma). It seems to us that these intervals are much too large when one considers both the shortness of the reaction times and the smallness of the differences between the reaction times of the different categories of judgment.

In order to overcome these two objections to the former studies, the present experiment employed a Dunlap Chronoscope which measured time intervals of only 1/1,200 second. Three observers were employed who gave a total of 7,000 judgment times on which the study is based.

The method employed utilized the lifted weight procedure described by Fernberger [5] for the method of constant stimuli. The method of obtaining the reaction times has been described by Fernberger and Irwin.[6] By this method the chronoscope is actuated by the lifting of the comparison weight and stopped by the observer pressing a pneumatic key when the judgment is verbalized. The only difference between the arrangement for the present experiment and that of Fernberger and Irwin was the placing of a standard weight before each comparison weight. The standard weights were 100 grams with 7 comparison weights ranging from 84–108 grams in steps of 4 grams. The space error was eliminated by using a turning top table and the time error was in the first order (standard weight always lifted first). The definition of the different categories

[5] S. W. Fernberger, On the relation of the methods of just perceptible differences and constant stimuli, *Psychol. Monog.*, 1913, No. 61, *cf.* esp. 10–19.

[6] Fernberger and Irwin, *op. cit.*, *cf.* esp. 507f.

of judgment were the same as those used by Fernberger.[7] The three categories of 'lighter,' 'equal' and 'heavier' were used.

The authors acted both as experimenters and as observers, alternating in the three functions of forming judgments on the lifted weight experience, presenting the stimuli and reading the chronoscope. Obs. C and D had some previous experience in the formation of psychophysical judgments and Ob. P had acted as observer in the previous time measurement experiments with the method of single stimuli. From each of Obs. C and D were obtained 7 series of 350 judgments and from Ob. P were obtained 6 similar series.

The results will be found in Tables I–III. In each table will be found the results for a single Ob. All tables are similarly constructed. There are three series of four columns each, one series for each of the three categories of judgment. In each series of columns will be found: 1, the number of cases; 2, the mean reaction times in sigma; 3, the sigma of the distribution and 4, the sigma of the mean. The results for each fractionated series of judgments will be found in successive rows in the order in which they were taken. In the bottom row will be found the averages for the entire series for each observer.

It is at once evident that the differences in mean reaction

TABLE I

REACTION TIMES, OB. C

Series	Lighter				Equal				Heavier			
	No. of Cases	Mean	σ dis.	σ mean	No. of Cases	Mean	σ dis.	σ mean	No. of Cases	Mean	σ dis.	σ mean
1	132	395	222	149.0	108	351	168	112.0	110	385	186	125.0
2	125	274	85	5.1	97	280	82	5.6	128	312	85	5.1
3	135	214	58	5.1	83	213	65	7.3	132	221	76	6.6
4	141	192	60	5.0	80	197	49	5.4	129	181	27	2.4
5	134	180	45	2.6	65	185	40	3.4	149	178	37	2.1
6	151	147	32	2.6	67	148	37	4.6	132	137	27	2.4
7	139	170	43	2.4	63	157	38	3.2	149	169	36	2.0
Ave.	137	225	78	24.5	80	219	68	20.2	133	226	68	20.8

[7] S. W. Fernberger, An introspective analysis of the process of comparing, *Psychol. Monog.*, 1919, No. 117, *cf.* esp. 13.

TABLE II

REACTION TIMES, OB. D

Series	Lighter				Equal				Heavier			
	No. of Cases	Mean	σ dis.	σ mean	No. of Cases	Mean	σ dis.	σ mean	No. of Cases	Mean	σ dis.	σ mean
1	136	348	106	9.1	69	324	96	11.6	145	317	96	7.9
2	162	289	77	6.1	39	299	68	10.9	149	266	76	6.2
3	165	289	68	5.3	54	284	62	8.4	131	293	74	6.5
4	160	264	64	5.1	65	243	69	8.5	125	243	55	4.9
5	175	181	38	2.8	49	184	52	7.4	126	182	39	3.5
6	171	175	29	2.2	49	184	28	4.1	130	180	31	2.7
7	165	139	28	2.2	62	140	29	3.6	123	143	33	3.0
Ave.	162	241	59	4.7	55	237	58	7.8	133	232	58	5.0

TABLE III

REACTION TIMES, OB. P

Series	Lighter				Equal				Heavier			
	No. of Cases	Mean	σ dis.	σ mean	No. of Cases	Mean	σ dis.	σ mean	No. of Cases	Mean	σ dis.	σ mean
1	110	458	87	5.8	110	502	134	8.6	130	424	60	5.2
2	119	252	72	7.3	93	248	75	7.8	138	235	65	5.5
3	106	218	41	4.0	112	211	37	3.5	132	212	43	3.7
4	147	170	35	3.0	75	161	31	3.5	128	168	37	2.5
5	119	186	50	3.4	44	174	40	4.1	187	166	23	1.4
6	114	152	45	4.2	93	157	46	4.8	143	149	46	3.8
Ave.	119	239	55	4.6	88	242	60	5.4	143	226	46	3.7

times are very small in every case and for every observer. These average differences will be found in Table IV. In the case of Ob. C, the average reaction times for the 'equal' judgments is *shortest;* for Ob. D intermediate and for Ob. P. longest. But these differences are so slight that when con-

TABLE IV

DIFFERENCES IN REACTION TIMES FOR THE DIFFERENT CATEGORIES OF JUDGMENT

Ob.	Lighter Heavier	Lighter Equal	Heavier Equal
C	1	6	7
D	9	4	5
P	13	3	16

sidered in relation with either the sigma of the distribution or the sigma of the mean, they are totally without statistical significance. The picture which they present is one of similar distribution affected by accidental errors.

And similarly when one considers the differences for the individual series they are in every case without statistical significance, although they are frequently much larger than the average differences. The largest difference is found for Ob. P. Series 1, Heavier-Equal. But it will be observed that when the differences are greater, the sigma of the distribution and of the mean also increase in magnitude. One may safely conclude that these results are in complete agreement with those of Fernberger and Irwin, indicating that there is no statistically significant differences in the length of the reaction times required for the formation of judgments of 'less,' 'equal' or 'greater' for the method of constant stimuli. And the results of the present study are in complete disagreement with those of Martin and Müller, George and Kellogg, who all found that the 'equal' judgments were formulated slower than either of the judgments of difference.

How may the difference between the present results and those of the three previous investigations be explained? It would seem doubtful if the present improved reaction time technique or the larger number of observations of the present experiment would account for this difference. It would seem more likely that the difference in the attitude of the observers toward the equality judgment would more adequately account for the differences in results. In both George's and Kellogg's experiments, it would seem that there was set up an attitude against giving equality judgments. This would be accomplished in the case of George by adding a large number of 'doubtful' and 'guess' categories and, in the case of Kellogg, by introducing a 'two-category' series in which equal judgments were eliminated by experimental direction. All observers in the present experiment had been previously trained in an attitude in which the equal judgment was accepted as a category equally valid and quite on a par with either of the categories of difference judgments. That

the attitude of the observer toward the equality judgment has a profound influence on the relative number of such judgments which will be given in an investigation has been several times demonstrated by Fernberger.[8] It seems reasonable that this same attitude toward the equality judgment would likewise have a profound effect on the judgment times as well.

Conclusion.—On the basis of a larger number of judgments and with an improved reaction time technique, we find that the judgment times for the three categories of judgment for the method of constant stimuli of 'less,' 'equal' and 'greater' are exceedingly similar and, in no case, are the differences great enough to be statistically significant.

[8] S. W. Fernberger, The effect of the attitude of the subject upon the measure of sensitivity, *Amer. J. Psychol.*, 1914, **25**, 538–543; The use of equality judgments in psychophysical procedures, *Psychol. Rev.*, 1930, **37**, 107–112; Instructions and the psychophysical limen, *Amer. J. Psychol.*, 1931, **43**, 361–376; Die Ungewissheitsurteile in der Psychophysik, *Arch. f. d. ges. Psychol.*, 1931, **80**, 273–290.

(Manuscript received September 10, 1933)

ANALYSIS OF MUSICAL APPRECIATION BY MEANS OF THE PSYCHOGALVANIC REFLEX TECHNIQUE [1]

BY MARY L. PHARES

Introduction.—Analysis of musical appreciation has been made largely by way of interpreting introspective accounts. There are no reliable objective techniques which tell adequately how the individual is affected by different types of musical compositions. Several investigators have studied the effects of musical stimuli upon metabolism, muscular energy, blood-pressure, respiration, electro-cardiograms, thresholds for sensory stimuli, and the knee-jerk, but the stimuli consisted chiefly of elementary musical sounds such as isolated tones and chords, short scales, bells and whistles. In very few experiments were musical selections used. Previous to the present investigation, the writer could find but one experimental study of musical appreciation by means of the psychogalvanic technique; that report [2] does not give any account of the effects on the psychogalvanic response to different types of music, nor of the effects of music, galvanically, upon different observers. It was with a view of determining the usefulness of galvanic technique as a method of studying musical appreciation that the present study was undertaken.

Apparatus.—The galvanometer, Wheatstone bridge, and form of electrical circuit were those used and described by Bayley [3] and by Patterson [4] except the source of potential, in the present study, was supplied by a two-volt standard storage battery instead of a dry cell. *E* made improvements in Bayley's and Patterson's recording

---

[1] This experiment was performed in the Psychological Laboratory of the State University of Iowa. Acknowledgment is given Dean C. E. Seashore for suggestions in completing the study.

[2] Diserens, Charles M. *The Influence of Music on Behavior*, 1926. A good review of the literature on musical appreciation is found in this reference.

[3] Bayley, Nancy. Study of Fear. Ph.D. Thesis, University of Iowa, 1926.

[4] Patterson, E. A qualitative and quantitative study of the emotion of surprise, *Psych. Monograph*, 40, No. 1, Univ. of Iowa Studies, 13, 1930.

apparatus by (1) exchanging their fan motor and speed-reducing arrangement of warped wooden wheels and cord belts for a 1/10 h.p. electric motor, (2) designing a metal worm-gear attachment thereby providing uniform movement of the kymograph drum at the rate of one revolution, using about 18 in. of paper, in four minutes, the maximum time of any phonograph record used, and (3) exchanging the siphon pens [5] for self-filling fountain pens,[6] four of which were attached to magnets [7] arranged by *E* in the following manner: The small arm of the magnet was replaced with a larger, firmer one at whose free end was soldered a cone-shaped holder 3.5 cm long, with its outer side left open for adjusting to the shape of the pen-barrel whose curved base was filed off and inserted freely and firmly into the metal holder attached to the magnet. This arrangement gave very uniform, clear ink-lines, and required no further attention after the pens were set into position. These five lines were simultaneously recorded: (1) One representing direction and extent of the galvanic deflection, traced by *E* with a pen in a holder and guide made to slide along a rod parallel to the scale; (2) A time line; (3) *O*'s signal line; (4) *E*'s signal line used for various notations such as giving a reference when the music started, indicating places of interruptions and other items essential for interpreting results; (5) A line traced by a pen [8] interrupted at each revolution of the turn-table of the phonograph in the outer room.[9]

In the first twelve observations, we used the electrode described by Ruckmick and Patterson; [10] in all others, we used the form designed by *E* as shown in Fig. 1, which provided the following improvements over the one first employed: Automatic control and comfortable position of *O*'s fingers; even resting position of the electrodes; constant contact with wire leads to the galvanometer; adjustable clamping of the electrode container to the experimental chair for accommodating different lengths of *O*s' arms.[11]

Victor's orthophonic large sized cabinet victrola, No. VV 8–12, 11130, was used. The records were Victor, Columbia, and Brunswick. Soft steel and fibre needles were used, according to the type which best brought out tones suitable for the size of room and distance of *O* from the machine.

Observers.—Six women and eight men, ranging in age from 22 to 45, observed. In music, eight had technical training; three, some, and five, intensive training. Seven played the piano, only; four, the piano and one other instrument. In psychology, all but three had training beyond the elementary course; six, in advance graduate units.

Materials and Procedure.—Series *A* was planned to secure psychogalvanic responses with supplementray accounts to music in four moods designated respectively as *gay*,

---

[5] Employed and described by Patterson, *Op. cit.*, 90.

[6] Made by the Inkograph Company.

[7] Made by the Howard Apparatus Company.

[8] Used only long enough to obtain a scale for locating in the phonograph records those points corresponding temporally with certain ones in the galvanic graph. It was not continued in observations because the mercury contact with the turn-table slightly affected the pitch of the music during sustained notes.

[9] The galvanometer, bridge, and recording apparatus were set up in a small room off from the larger one occupied by *O* and the phonograph.

[10] Ruckmick, C. A., and Patterson, E. A simple non-polarizing electrode. *Am. Jr. Psych.*, 1929, 41, 120–21.

[11] To guard against oily sediment collecting in the chemical, and heating of the water, duplicate sets of this electrode, Fig. 1, were in readiness, and a fresh clean pair was employed at the beginning of each observation-hour.

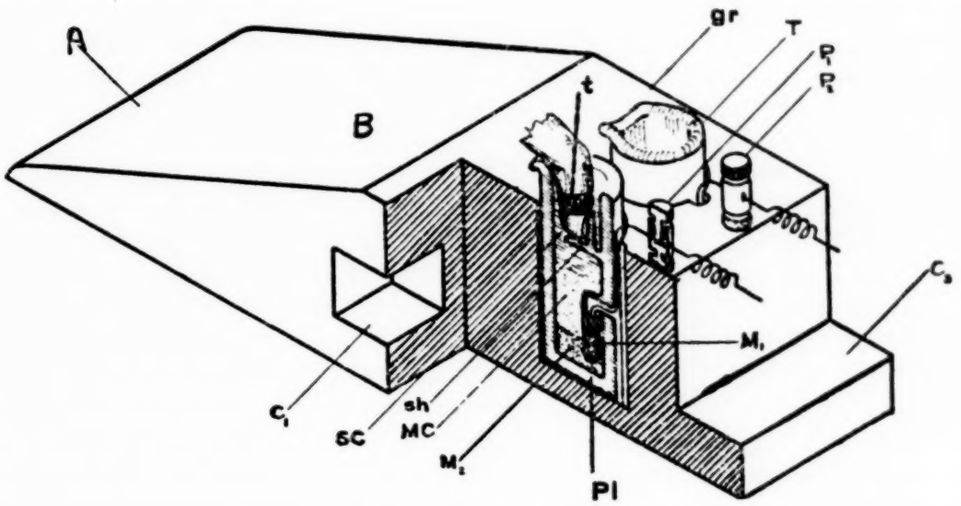


FIG. 1. Electrodes in block container.

B, Block container
A, Arm rest
$C_1C_2C_3$, Points for clamp (C_3, opposite of C_1)
P_1P_2, Posts for wire leads to galvanometer
t, Tape placed at first crease of finger
T, Tube in which finger is inserted (11.5 cm high, 4 cm in diameter)
gr, Groove in large tube
sh, Shelf on which finger-tip rests (1.3 cm in diameter, and 4.8 cm below upper edge of tube)
SC, Sodium Chloride solution
MC, Mercurous Chloride
M_1, Mercury in small inner tube
M_2, Mercury in large tube
Pl, Platinum contact between tubes of mercury.

melancholy, *triumphant*, and *tragic*. For each mood, eight compositions on phonograph records were selected. During each of eight weekly observations made by *O*, *E* presented four different selections, playing as much from each as one face of a record contained. Following are the selections [12] in the order presented.

As soon as *O* entered the room, *E* directed him to sit in a large morris chair and rest for about five minutes. *E* then prepared *O*'s fingers for the electrodes by (*a*) cleansing them with cotton moistened with alcohol, and (*b*) wrapping with Tirro Waterproof tape placed at the crease of the first joint (see Fig. 1), for providing uniform surface contact in the electrodes during the period and succeeding observations. *E* then gave *O* a page of typewritten directions which read as follows:

"Arrange yourself comfortably in the chair; hold your attention upon the music from its beginning. Assume the attitude of sitting in an audience listening to the productions by a famous orchestra. You will hear four selections of music; after each,

[12] The figure in parentheses following the name of each selection indicates the number of *O*s who gave their highest rank to that number. Abbreviations used in the list are: Pt. for Part; Mv. for Movement; Sym. for Symphony; and Orch. for Orchestra.

	Name of Selection	*Composer*	*Performers*
(0)	Capricieuse	Elgar	Heifetz, with piano
(5)	Carneval Overture, Pt. 1	Dvorak	Chicago Sym. Orch.
(2)	Coppelia Ballet	Delibes	San Francisco Sym. Orch.
(3)	Jupiter (Planets)	Holst	London Sym. Orch.
(7)	Quintet in A Major, 4th Mv., Pt. 6.	Schubert	Forellen Quintet
(9)	Rosamunde Overture, Pt. 1	Schubert	Harty & Halle Sym. Orch.
(2)	Sym. No. 8, in F, Pt. 3	Beethoven	Royal Philharmonic Orch.
(0)	Syncopation	Kreisler	Kreisler, with cello-piano
(0)	Quartet in E-Flat, 2nd Mv., Pt. 3.	Beethoven	Lener String Quartet
(0)	Quartet in G-Minor, 1st Mv.	Debussy	New York String Quartet
(12)	Sym. No. 5 in E-Minor, Pt. 1, 1st Mv.	Tschaikowsky	Chicago Sym. Orch.
(4)	Sym. No. 6, Pt. 2, 1st Mv.	Tschaikowsky	New Queen's Hall Orch.
(0)	Quartet, A-Minor, Pt. 0, No. 2	Brahms	Lener String Quartet
(0)	Eroica, 2nd Mv., Pt. 4	Beethoven	London Sym. Orch.
(0)	Iberia, No. 2, Pt. 3	Debussy	Royal Philharmonic Orch.
(12)	Unfinished Sym., 2nd Mv., Pt. 4.	Schubert	New Queen's Hall Orch.
(1)	Sym. in D-Minor, 3rd Mv., Pt. 8.	Franck	New Queen's Hall Orch.
(1)	Sym. No. 1, 4th Mv., Pt. 8	Brahms	London Sym. Orch.
(10)	Marche Militaire	Schubert	Minneapolis Sym. Orch.
(0)	The 1812 Overture (Finale)	Tschaikowsky	New Queen's Hall Orch.
(6)	In Springtime Overture, Pt. I	Goldmark	Chicago Sym. Orch.
(5)	Rigoletto, Introduction	Verdi	Creatore and his Band
(4)	Stars and Stripes Forever March	Sousa	Arthur Pryor's Band
(1)	Soldier's Chorus March	Gounod	Arthur Pryor's Band
(0)	Moonlight Sonata, No. 2	Beethoven	Paderewsky, piano only
(3)	Lucia-Sextette, Act 2	Donizetti	Creatore and his Band
(13)	Death of Boris, Act 4	Moussorgsky	Chaliapin, with Orch.
(0)	Sym. No. 6, Pt. 8 (Finale)	Tschaikowsky	New Queen's Hall Orch.
(4)	Rigoletto—Quartet	Verdi	Creatore and his Band
(0)	Prelude in C-Sharp Minor	Rachmaninoff	Victor Concert Orch.
(6)	Tod Und Verklaerung, Pt. 1	Strauss	Royal Philharmonic Orch.
(2)	Trovatore—Miserere	Verdi	Creatore and his Band

give a report. Analyze quite definitely how you respond to the music, and answer these questions:

1. What mood or general feeling-tone does the music induce in me?
2. What are the specific factors of the music that I do, or do not enjoy?
3. Was my attention distracted? Why? During what part of the music?
4. Did I become more relaxed or more tensed?
5. With a scale of +4 to −4, rank each selection with reference to the degree of feeling-tone induced; thus:

+4 to indicate *very pleasant*
0 to indicate *indifference*
−4 to indicate *very unpleasant.*

6. Tap the signal-key at your right hand:
   Once when definitely aware of pleasure
   Twice when definitely aware of displeasure
   Three times when aware that attention had been off the music.[13]

[13] This could not be carried out because it diverted attention from the music and changed the resistance of some *O*s.

When the music and galvanic response finished, *E* released *O* from the electrodes, and wrote his answers to the questions. In addition, *E* without giving suggestions, always endeavored to obtain data upon sensory experiences, associations, and imagery. Because *O* inveriably gave a larger deflection during his initial observation, *E* repeated the first selection in each mood-group during a subsequent period.

Precaution was taken against having an individual observe when he was not in his usual mental and physical health and comfort. Except for a few 'make-up' periods with two *O*s, the time of day for a given subject remained the same throughout the experiment, which lasted from November 1, 1928 to June 20, 1929.

Series B was planned to determine whether or not (*a*) form or size of galvanic curves could be used to indicate qualitative differences in *O*s' responses to music; (*b*) increasing familiarity with the music would affect size or form of the galvanic curve. From Series *A*, for each *O*, *E* obtained two sets, each containing four selections to which *O* had given his highest ranks—one for each of the moods, *gay*, *melancholy*, *triumphant*, *tragic*, as indicated by the *O*. In that order, a set was played three times in succession during each of three 50-minute periods—Set I in the first, third, and fifth, and Set II in the second, fourth, and sixth weeks. For the first round during a period, one face of a phonograph record was played. For the second and third rounds, only that part most enjoyed by *O* was used, unless such was the latter part of the record; in that case, *E* played the entire face because beginning in the middle of a record annoyed *O*. After each repetition, an account was taken only with reference to (*a*) degree and quality of affective experience, (*b*) distraction of attention, and (*c*) relaxation or tension and alertness in *O*. Other points of importance to the problem, as offered by *O*, were noted, however. For the purpose of establishing an optimum mental set, *E* began the observation hour with a short musical selection not included in the sets above. No galvanic readings were taken for these introductory numbers.

Method of Calculating the Value of Deflections.—Two of *O*'s psychogalvanic curves, each 4 cm in length, are not of equal value, or comparable, if his skin resistance to the electric current during one was 12,000, and the other, 16,000 ohms; they are comparable only if his resistance taken immediately preceding each deflection gave the same values. Since *O*'s resistance varied from time to time and from other *O*s, the galvanic results were made comparable by converting them into uniform measures. This was done by determining the absolute change in ohms for each deflection from the two values, (*a*) *O*'s initial resistance, and (*b*) length of graph representing his galvanic response to the stimulus, through the method of direct calibration of the apparatus with a variable resistance box substituted for *O* and the electrode contacts. *E* then calculated the percentage which the absolute change was of the initial resistance. Experimentation [14] shows that since *O*'s initial resistance is a factor in conditioning the size of the deflection, the percentage change gives a more accurate value of *O*'s response than will the absolute change.[15]

Treatment of the Galvanic Scores. A. Mood-grouping.—*O*'s report on quality of mood experienced did not correspond in all cases with *E*'s tentative mood-classification of selections. All observations[16] during which *O* reported experiencing a mood-quality

[14] Lauer, Alvhh, R. Reliability of the galvanic reflex. *Am. Jr. Psych.*, 1929, 41, 263–270.

[15] Throughout the present study, all numbers representing values of galvanic responses, such as single deflections, means, and ranges, are expressed in percentage changes in resistance by means of the abbreviation, '*R-changes*.'

[16] 'Observation,' as used here, means listening to one selection of music. In Series *A*, either four or five selections were played during each observation hour; in Series *B*, parts of four different selections, each three times, were presented during the hour.

of *gay* were placed in one group; of *melancholy*, in another; of *triumphant*, in a third; of *tragic*, in a fourth; of any *other* type, in a fifth, designated as '*Other moods*' group; all observations of *non-mood* experiences were placed in a sixth group.[17]

B. Computations.—The writer found no directly applicable reliability formula for testing differences between the galvanic scores since they are relative measures, but applied those available statistical techniques which best fit the conditions, and computed the following values:

1. *The percentage frequencies* of (*a*) affective and non-affective responses to the initial presentations of the 36 selections, (*b*) the total number of zero R-changes, (*c*) 'indifference' and 'dislikes,' (*d*) mention of musical elements in reports corresponding with R-changes *above* and *below O*'s mean.

2. *Ranges and means* of R-changes of (*a*) each selection of music, (*b*) each observer, (*c*) each mood-division for *O*s separately and collectively, (*d*) increased and decreased R-changes separately.

3. *Correlations*[18] between *O*s' scores; and *standard errors*[19] of obtained differences between means of (*a*) increased and decreased R-changes, (*b*) individual selections, (*c*) different mood-divisions, (*d*) re-presentations of selections.

From all of the foregoing computations, comparisons in the *amount* of galvanic responses were made. To ascertain further the usefulness of galvanic technique for discriminating æsthetic experiences from music, *form* of the graph representing *O*'s change in resistance was analyzed.

Comparisons and Interpretations. A. Selections. 1. *General Factors.*—With reference to the type of dominant affective quality induced by a given selection, there was variation among *O*s for each number except that selection used from 'Death of Boris' which elicited the tragic mood in all subjects. Initial presentation of the 32 selections gave the following distribution of the 504 observations with reference to the frequency of affective experiences:

Tragic,	6 percent	Gay,	18 percent
Triumphant,	8 percent	Other moods,	28 percent
Melancholy,	11 percent	Non-moods,	29 percent.

These differences may be due in part to (*a*) the narrow range of selections used, and (*b*) *O*'s preformed associations and experiences with music, as indicated by the fact that some *O*s experienced *gay* feeling from selections which induced the

[17] These numbered 144 and do not include 9 cases of zero ranks in verbal reports of the following nature: "I did not experience any mood other than 'dislike' for the selection—no specific mood-quality." All single 'dislikes' were regarded as affective responses and included in the group of 'Other moods.'

[18] For the correlations, we used the formula in Garrett, H. E., *Statistics in Psychology and Education*, 1926, 168.

[19] For standard errors, we used the formulas for calculating *t* values given by Fisher, R. A., *Statistical Methods for Research Workers*, 1928, 2nd ed.

triumphant in others; *melancholy*, from selections which induced 'serene,' 'peaceful' attitudes in others; *tragic*, from selections for which other *O*s reported "No affective response, design too complicated for enjoyment," "Melody and composition did not appeal to me, did not attend to it long, no mood experienced."

TABLE I

SHOWING RESULTS FOR EACH SELECTION OF SERIES *A* [20]

1	2	3	4					5	6	7		
Selections	Total Mean R-ch.	Musical Struct.	Freq. of Mood Classification					Freq. No Mood	Freq. Dislike	Freq. of R-ch.		
			G	M	Tp	Tg	O			Zero	Inc.	Dec.
1. Unf S.....	24.2	1	—	12	—	—	2	0	0	0	14	0
2. Bor.....	17.5	1	—	—	—	14	—	0	1	1	0	13
3. Ts S 5.....	14.1	1	—	13	—	—	1	0	0	0	8	6
4. Rig. I.....	12.8	1	5	—	8	—	1	0	0	0	0	14
5. Path 2.....	11.6	1	—	9	—	—	5	0	0	4	6	4
6. Rig. Q.....	11.3	1	—	1	—	4	8	1	0	3	0	11
7. Capr (1).....	11.1	1	11	—	—	—	3	0	0	0	1	13
8. Sp O.....	10.4	2	7	—	6	—	—	1	0	1	0	13
9. M M.....	10.4	2	—	—	13	—	1	0	0	1	0	13
10. Syn.....	9.6	1	10	—	—	—	3	1	0	3	9	2
11. Trov.....	8.6	2	—	3	—	5	5	1	0	3	0	11
12. Ros O.....	8.1	1	11	—	—	—	1	2	0	3	0	11
13. Eroi.....	8.0	2	—	2	—	—	11	1	2	2	0	12
14. Capr (2).....	7.9	1	8	—	—	—	6	0	0	5	8	1
15. Car O.....	7.7	2	10	—	—	—	2	2	0	2	0	12
16. Quin.....	7.4	2	8	—	—	—	3	3	1	4	2	8
17. Sext.....	7.0	3u	—	4	—	3	5	2	3	5	1	8
18. Star.....	6.6	2	1	—	9	—	1	3	0	5	0	9
19. Path 8.....	5.7	2	—	5	—	—	5	4	0	6	2	6
20. Cop B.....	5.2	2	8	—	—	—	4	2	1	4	0	10
21. Iber.....	5.1	2	—	3	—	—	8	3	2	5	0	9
22. Tod.....	4.5	3c	—	1	—	6	4	3	4	9	0	5
23. M S (2).....	4.4	2	—	—	—	—	7	7	0	8	5	1
24. Fr S (1).....	3.1	3c	1	—	1	—	4	8	1	10	0	4
25. Jup.....	2.9	2	4	—	—	—	7	3	3	10	1	3
26. M S (1).....	2.9	2	—	1	—	—	7	6	0	9	1	4
27. Sol C.....	1.8	3u	2	—	3	—	5	4	4	11	0	3
28. Be Q (1).....	1.5	3c	—	2	—	—	2	10	0	11	0	3
29. Be S.....	0.08	3c	2	—	—	—	4	8	1	12	0	2
30. Br S 8.....	0.04	3c	—	—	1	—	3	10	2	13	0	1
31. Sole.....	0.03	3c	—	—	1	—	10	3	10	13	0	1
32. Be Q (2).....	0.00	3c	—	—	—	—	1	13	1	14	0	0
33. Fr S (2).....	0.00	3c	—	—	1	—	1	12	1	14	0	0
34. De Q.....	0.00	3c	—	—	—	—	4	10	3	14	0	0
35. Br Q.....	0.00	3c	—	—	—	—	4	10	3	14	0	0
36. Pr R.....	0.00	3c	—	—	—	—	3	11	3	14	0	0
			88	56	43	32	141	144	46	233	58	213

[20] Explanations for Table I: *Section* 1: For full name of the selection, refer to page 122. The number in parentheses following a name indicates the first or the second time that selection was played in Series *A*. See full explanation, page 123.

Section 2: The mean of all *O*s' scores, including the zero galvanic readings.

Section 3: Verbal reports revealed that musical structure of a selection, as it appeared to the majority of *O*s, could be characterized with a general descriptive phrase pertaining to the positive-negative mood attitude, as indicated below:

1 indicates *satisfying.*

2 indicates *fairly satisfying.*

3c indicates *not satisfying* because the selection was *too complex.*

3u indicates *not satisfying* because the selection was *too undeveloped.*

Section 4: G refers to gay; M, melancholy; Tp, triumphant; Tg, tragic; O, 'other moods.' The latter include 28 'dislikes' during selections inducing no specific mood-quality.

Section 5: These values refer to the number of *O*s reporting 'no mood felt' for the selection.

Section 6: These values refer to the number of *O*s reporting 'dislike' for the selection.

Section 7: Inc. refers to increased R-changes; Dec., to decreased R-changes.

As shown in Table I, there were 28 selections for which some one or more of the *O*s reported 'No mood experienced'; there were no selections for which all *O*s gave that report. All non-mood reports were on selections which also showed zero galvanic readings. There were four selections which showed no zero galvanic readings. Comparison of the musical numbers in terms of size of *O*s' mean galvanic responses shows considerable variation from selection to selection.

Supplementary reports show that those numbers which were satisfying and appealing for their general musical structure gave the highest R-changes, were seldom disliked, and rarely corresponded with non-mood reports. Selections which appeared to *O* either too complex or too undeveloped in their musical structure gave either zero galvanic readings or small mean R-changes.

From the total number of 1509 observations (Series *A* and *B*), there were 120 reports of *indifference*, and 46 of *dislike*. Sixty-nine percent of the 'indifference' and sixty-one percent of the 'dislikes' were reported for selections inducing no definite mood-quality. Twenty-six percent of the 'indifference' and thirty percent of the 'dislikes' were for selections inducing qualities in the 'Other moods' group. Verbal reports indicate that *O*s became indifferent to, or disliked, selections not appealing to them emotionally. *O*s' reasons for their attitudes of indifference and dislikes are listed below in order of greatest frequency:

a. Design too complicated. Don't understand that kind of music; don't enjoy music which makes me labor too hard to get anything out of it.
b. Unfamiliarity with the selection.
c. Mechanical defects in the production.
d. Over-familiarity with the selection.
e. Design uninteresting—lack of development of melodic and harmonic structure, tone color, tempo, or intensity.
f. Unfavorable associations previously established for the selection.
g. Not in the mood for the type of music.
h. Mental set was not yet ready for the music.
i. Distractions (Noises outside, desire to change position, etc.).

Of these reasons, a tabulation was made separately for the *musically trained* and the *untrained*. Reasons given most often by the former, in order of frequency, are (*c*), (*e*), (*g*); by the latter, (*a*), (*b*), (*c*).

2. *Musical Elements.*—Verbal reports indicate that when the design of a selection was very inconspicuous, *O* was quite aware of this musical element being an important factor in his æsthetic appreciation; but when the selection had a marked pattern or thematic structure running through it, *O* was not so conscious of that as an element, apart from the melodic or harmonic aspects. Thus, regarding Tod Und Verklaerung, a frequently reported datum was:

"The selection is so formless—no pattern running through it that catches and holds my attention; that is disturbing to me. I can't get much out of it."

When the selection possessed a complex network of design, *O* was often attentive to this characteristic as conducing to his lack of enjoying the selection. Thus, for Bach's Fugue which was used as an introductory number only, a very frequently reported datum was:

"It seemed so mixed up; about as soon as I got hold of one little pattern then another rushed in upon it,—and I lost the former,—that was annoying to me."

Toward the complexity of design, it was interesting to note the individual differences in attitude among *O*s. A frequent remark from musically trained students about the selection possessing intricate design was:

"I am sure that if I were to hear that several times again, I would perceive more in it to appreciate; it is too complex, too intricate to comprehend from one hearing."

Whereas, students not musically trained remarked:

"I don't enjoy that kind of music; it forces my attention and that is annoying to me. I lose interest, and my mind wanders to other things. I don't appreciate music by way of making labor and hard study out of it."

Frequency of reports of musical elements above and below *O*s' total mean R-changes were tabulated separately for displeasurable and pleasurable effects. Selections disliked because of (*a*) melody, (*b*) harmony, or (*c*) design, elicited respectively 2.0, 2.8, and 2.5 times as many smaller deflections as larger ones. As factors inducing displeasure, design was mentioned 5 times, and harmony 1.5 times, as often as melody. There was no selection whose rhythm or tone color was displeasing to *O*. Observations during which *O* was especially cognizant of (*a*) melody, (*b*) tone color, (*c*) harmony, or (*d*) rhythm as a factor in his musical enjoyment, corresponded respectively 13, 7, 4, and 1.5 times as often with the larger as with the smaller deflections. Observations during which *O* was especially cognizant of design as a factor in his musical enjoyment corresponded 2.4 times as often with the smaller as with the larger deflections.

B. Observers. 1. *Ranges.*—The largest R-changes range from 21 for H to 49 for D, and represent either a melancholy or a tragic selection for each observer but H whose largest included both triumphant and tragic numbers.

2. *Means.*—The means calculated for each *O* separately range from 3 to 9 in the group of '*Other moods*'; 9 to 17 in the combined results of *gay*, *melancholy*, *triumphant*, and *tragic;* and 5 to 10 for all his exposures (approximately 108 for each *O*).

C. Types of R-changes. 1. *Significance in Terms of the Subjective Report.—Zero galvanic readings* distributed fairly evenly among all *O*s, and corresponded with verbal reports which indicate analyzing attitudes, displeasure, indifference, inattention to the stimulus, random thoughts, mild relaxation, mild affective and non-affective responses.

Changes of increased resistance correspond with protocols indicating that *O*s experienced relaxation during the music. There were no reports of kinæsthetic sensations or 'urges' to movement and alertness during those deflections. They consistently correspond with such descriptions as 'very pleasing and soothing,' 'lulling melody,' 'full of restful beauty and charm,' 'subdued tones of particularly appealing har-

monic values,' 'unbroken, even beat in the rhythmic pattern,' 'slow, quieting, rambling movement in the general structure of the selection.' While the protocols show such terms as *gay*, *melancholy*, and others indicating affective-tone experienced from the 'relaxing' selections, in none of the descriptions is there a stress of the mood-experience; on the other hand, there is emphasis of such descriptions as 'peaceful,' 'dreamy,' 'soothing,' 'pensive,' 'reverent,' 'quieting,' 'serene,' 'dainty,' 'calm,' 'care-free,' 'playful.' These terms are more indicative of the general physiological adjustment of relaxation than of true emotional experiences. *O*'s verbal reports indicate that mood-quality was frequently secondary to his associations set off by the selection which was often the thing actually characterized affectively more than his own reactions. From all of the various factors involved, it appears that (*a*) the increased R-change, and extent of it, is sequential to increased relaxation induced by *O*'s concomitant perception and appreciation of many closely related stimuli-characteristics interwoven much as a gestalt—the affective quality being but one of the interrelated items in the configuration; (*b*) and that differences in amounts of increased R-change are due to a complex of factors about the selection.

Changes of decreased resistance, according to *O*s' reports, represent increased alertness, tension, stimulation, during the music. They correspond with such descriptions as: "Strong kinæsthetic sensations," "was moved to action," "I wanted to let go, sway, tap, beat or march with the music," "there was hope, courage, confidence, triumph, force, vigor, gaiety, patriotism expressed in the music," "I was alert, wanted to go, to move, to make progress in some way." All protocols in which *O* reported an active alert response to rhythm, tempo, sudden changes in intensity or forward movement of the selection correspond with decreased R-changes. Significant data for interpreting decreased R-changes are the facts that (*a*) all reports of 'displeasure' from the music, or from mechanical defects, correspond with either zero galvanic readings or decreased R-changes; and reports which indicate the strongest displeasure are those corresponding with de-

flections; (*b*) supplementary reports revealing strong analytic attitudes, with little affective response, correlate with either zero galvanic readings, or decreased R-changes; (*c*) the 16 deflections with verbal reports of, "Not certain that I made an affective response," place into two groups: (1) moderately active, analytic responses corresponding with decreased R-changes and (2) passive responses of slight relaxation corresponding with increased R-changes. In both the latter, the deflection is very small compared to those with verbal reports emphasizing cognizance of affective tone. Affective responses correspond with both types of R-changes, but the experience of 'organic sensations' as reported by *O*s correlate most frequently with decreased R-changes.

2. *Frequency and Size.*—Table II presents 4.1 times as many R-changes as zero readings, and 2.3 times as many decreased as of increased R-changes. As shown by (*a*) the

TABLE II

RANGES, MEANS, AD'S OF THE TYPES OF R-CHANGES

	Increased	Decreased	Zeros
* Cases	368	848	293
Range of R-changes	2–49	2–32	
Means	17	11	
AD's	9	4	
Limits of middle group (57.5 percent of the cases)	8–26	7–15	

* Includes the observations of Series *A* and *B;* three of the latter were not used, one during gay music because *O* moved his fingers in the electrodes, and two during melancholy music because of interruptions. Galvanic results of these interruptions are discussed in Section *D* of *Comparisons and Interpretations.*

AD's and (*b*) the ranges and means, respectively, there was greater variability of galvanic response, and there were larger R-changes made to the soothing, restful musical selections (increased resistance), than to the stimulating, arousing type of music (decreased resistance).

Correlation between the 14 *O*s' mean total increased and mean total decreased R-changes gave .68 ± .10; correlation between *O*s' largest increased and their largest decreased R-changes gave .69 ± .09. These comparisons indicate that

those *O*s who showed the larger increased R-changes tended to be the subjects who also gave the larger decreased R-changes.

Mood-groups were compared as to type of R-change; the results show that no one emotional quality, apart from other factors, was consistently and significantly connected with either increased or decreased R-changes.

Mood-groups were compared in the amount of R-change from the following levels: (*a*) By using selections showing equality of *O*s' ranks in musical structure. This restricted comparison, employed in lieu of using all observations in a given mood-division against all in another mood-group, is justified because a highly significant mean difference was obtained from the following comparison: (1) All observations showing ranks 1 and 2 in musical structure, vs. (2) all showing ranks 3c and 3u with means of 11 and 4, respectively. Consequently, the mean of all *gay* selections having rank one in musical structure was compared with the mean of all *tragic* selections having rank one, et cetera. Results from the eight comparisons available gave no significant differences between mood-divisions; from this, we may tentatively conclude that music written in one mood will elicit as large galvanic responses as music written in another mood, other stimuli characteristics being equal to *O* for the selections being so compared. (*b*) By using *O*s' scores from the first day's initial rounds of their selected musical numbers; results from the twelve comparisons available gave no significant differences between mood-groups. (*c*) By using the amount of decrement in galvanic response during repetition of the selected numbers. The mean of *O*s' scores for each round and day, show for each mood-group similarly, that the effect of representation within both the fifteen-minute and the two-week intervals was that of a general decrease, with few exceptions, in the mean scores. Reliability values of obtained differences between the groups, in the amount of decrement, were not significant; this indicates that lowering in the galvanic response, with successive repetitions of the selections, did not offer criteria for qualitatively differentiating moods experienced from music.

In ninety percent of the cases, R-changes above *O*'s mean corresponded with his verbal reports indicating the more intense affective experiences, regardless of quality of the emotion. Evidence on this point was gained from the second and third repetitions of selections during the same observation-hour; the amount of R-change lowered, and according to *O*'s report, his affective response fell off. Of the 293 zero galvanic readings, 49 percent corresponded with 'non-mood' experiences; the remainder, with reports indicating '*mild* affective' responses. Taking all the data into consideration, they appear to indicate that amount of R-change is of possible service in grossly differentiating the more intense from the less intense affective experiences.

D. Form of the Curve.—Figure 2 shows that form of the galvanic curve presented the three [21] general characteristics of *direction*, *width*,[22] and *height*. Width and direction were discussed in Section *C*, but some of the data here bear also upon direction. Height of the curve, as measured by the angle of deflection, varies with (*a*) amount of R-change (width of the curve), and (*b*) time [23] of occurrence of the change in resistance, as illustrated in Fig. 2. When curves of the same time but of different widths (*a*, *b*, *c*, Fig. 2), are compared, the widest ones, which are also those of highest rise, are indicative of greater tension and alertness, according to their corresponding protocols; when curves of the same width but of different time (*a'*, *b'*, *c'*, Fig. 2), are compared, those of shortest time, which are also those of highest rise, are indicative of greater tension and alertness; when curves of the same height but of different time and width (*c*, *c'*, Fig. 2), are compared, the one of shorter time (*c'*) indicates greater

[21] A fourth aspect, time of the return sweep, was not serviceable in this study because of using continuous stimuli and a critically-damped galvanometer.

[22] Width of the curve is dealt with only in terms of the amount of ohms change in resistance, since *O*'s initial resistance must be taken into account when comparing sizes of deflections.

[23] This characteristic could not be checked statistically because continuous stimuli were employed, and the nature of the stimuli, musical selections, did not afford enough variation to get a sufficiently large quantity of extreme contrasts in time; but the fortunate accidents discussed in this section gave pertinent clues on time. The criterion used was the period of the galvanometer—*i.e.*, R-changes not extending beyond the period of the galvanometer were designated 'rapid'; those beyond that period, 'slow.'

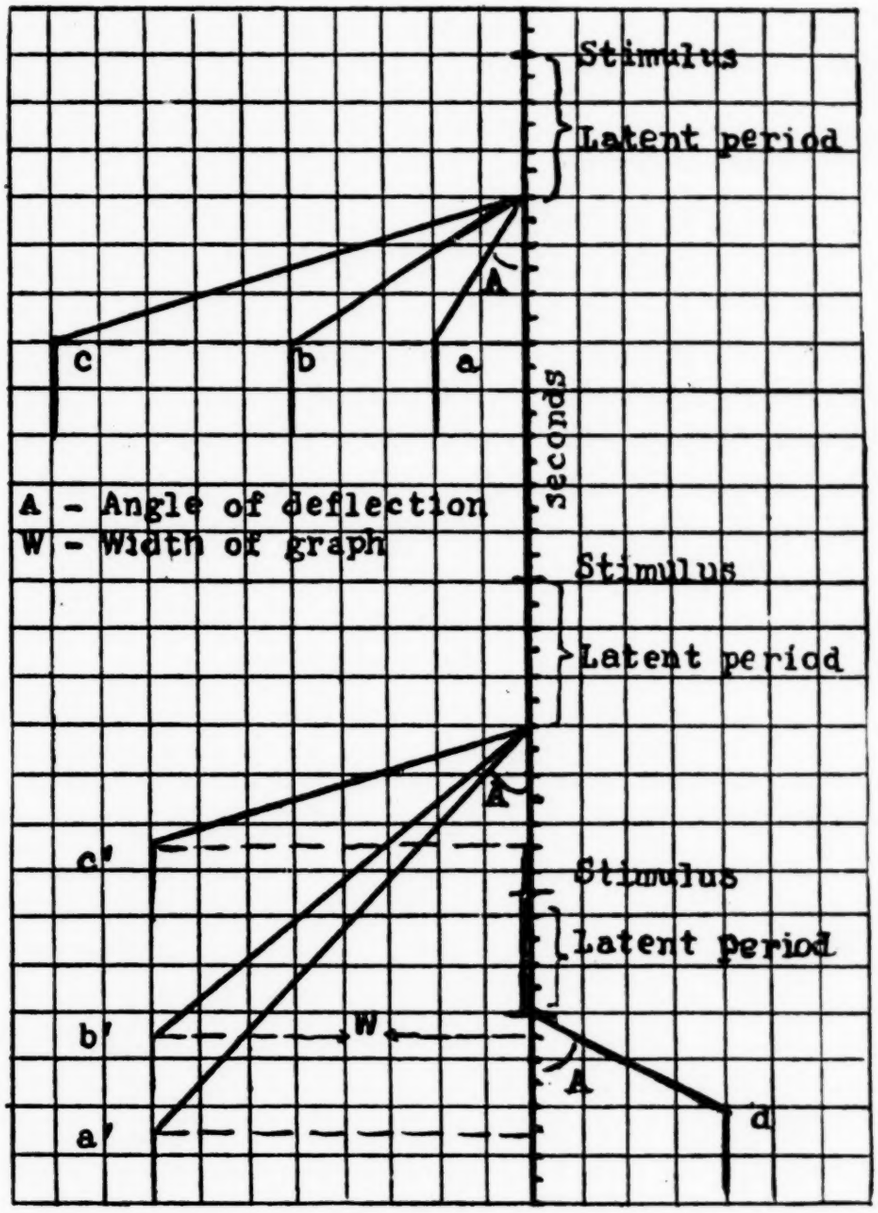


FIG. 2, showing the general characteristics of a galvanic graph.

tension and alertness. Hence, height of the curve, as well as direction, is coupled with the factor of alertness and tension.

Bearing upon this point are data from fortunate accidents during three selections. The results of one case are shown in Fig. 3. During Schubert's Unfinished Symphony, the fibre needle split. *E* allowed the noise to continue, since it vitiated *O*'s response to the music. *O* reported the following:

"I was thoroughly enjoying those soft, lulling tones—then when that scraping sound started—it was not a shrill penetrating scratch, and I didn't pay much attention to it

at first, except to feel that the music was spoiled,—I thought it was a click in the record, and the noise would cease in another round or two; but it continued and became much worse; it then was irritating and I grew tense,—I wanted to call to you to shut the machine off, but remembered that I should not move in the electrodes, and I grew more tense all the while it continued."

This curve shows a latency of about 20 sec. after the scraping noise started—then a change of decreased resistance corresponding with *O*'s experience of alertness. The rise in the deflection, *t*, coincident with increased 'tension,' is steeper than that at *r*, coincident with 'relaxation.'

Another interruption occurred as the result of some one outside the experimental room, who, inadvertently failing to read *E*'s sign, opened and slammed the door shut. About four sec. after this stimulus, the deflection reversed from increased to decreased R-change with a sudden rise in the curve, as shown in Fig. 4. *O*'s report was:

"From the dreamy, restful, calm feeling that I was enjoying during those soft, sombre spots in the music, I was suddenly aroused—I was almost startled; became tense and could not get back into the music. I experienced organic sensations of a sinking feeling inside, and a catching of my breath." Pathetique, 1st Mv., Pt. 2—Observer *N*.

Again there resulted a rise in the deflection, *t*, corresponding with *O*'s experience of 'tension,' steeper than that at *r*, corresponding with *O*'s experience of 'relaxation.'

The third accident occurred when observer *D* was listening to Schubert's Quintet for the first time. She always requested the shade drawn to rest her eyes while listening. After the music had played about $2\frac{1}{2}$ minutes, the deflection, suddenly quickened its course giving a high rise of increased resistance. *E* stepped out to examine connections; the phenomenon could have resulted if a wire snapped out the mercury contact, or if *O* withdrew her fingers to an appreciable extent. The physical conditions were intact. Within another instant, the shade spring gave way—the curtain springing to the top with a loud noise. *O* startled, then noticed *E* by her side for the first time. *E*, returning immediately to the galvanometer, found the deflection reversed and presenting a wide curve of decreased resistance. In the verbal report, *O* remarked:

"I fell into a doze soon after that selection started—nothing there to hold one musically—too light, fluffy, undeveloped; I confess I was totally gone and don't know how you got out here."

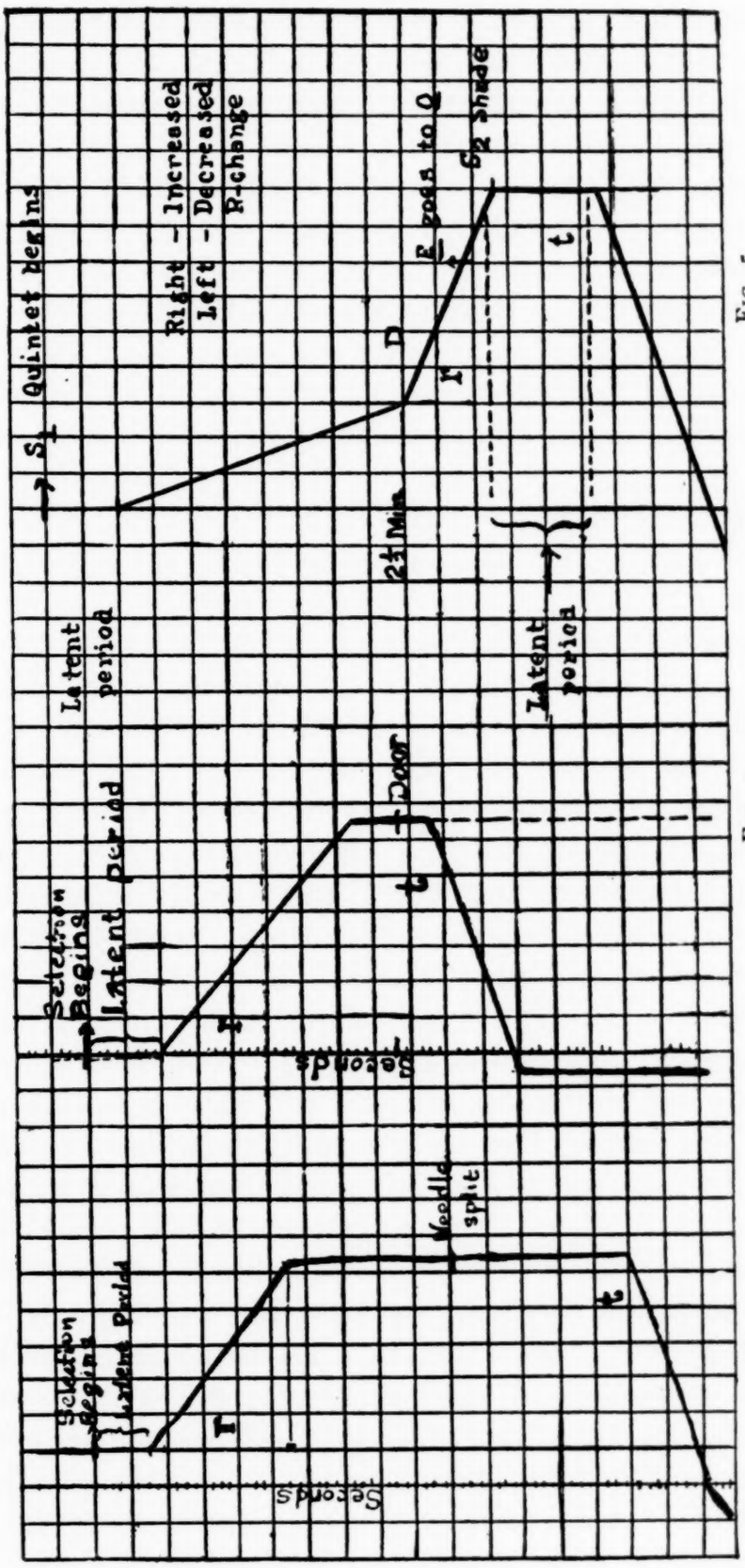


FIG. 3

FIG. 4

FIG. 5

Graphs of deflections elicited during interrupted observations of selections which were inducing relaxation in *O*. In the electrical circuit used in this experiment, a deflection turned to the left indicates decreased R-changes; one to the right, increased R-change.

O emphasized her experience of a complete let-down throughout her muscles after she began to relax as the music advanced. Figure 5 represents the galvanic response during the entire period *O* was in circuit. The line DS_2 corresponds with the time *O* fell 'totally gone' into the doze, a condition which makes the probability very small that she was also at that moment experiencing an affective quality, yet there is the sudden high rise in the curve at *D*. From S_2 on to the lower end of the graph represents the period from the time the unforeseen S_2 crowded in upon *O*'s mental and physical state, one of practically complete unreadiness for the imposing stimulus.

The protocol in this case revealed that *O* experienced elements of startle, irritation, surprise, wonder, chagrin, and laughter; but in addition, she was most certain of a sudden change from one highly adapted physiological state, relaxation, to a momentary extreme of the opposite, tension. To label the deflection of decreased R-change with one of these emotional qualities without analysis of the total situation would be giving an inadequate interpretation. In Fig. 5, we see that the angle *r*, at the point of deflection subsequent to *O*'s 'falling totally gone into the doze,' is equivalent to that at *t*, the point of deflection subsequent to S_2, yet the situation for the former elicited practically a non-affective response, while that for the latter, a strongly affective response. Hence, we see that when the total situation, each time, is taken into account, a given curve is found to correlate not with any one subjective factor, but with a complex which the total physical situation suddenly superimposes upon a large number of neurological configurations, prepotent and acquired; and notwithstanding the influence of affective qualities, or of temporal succession of them, we must not lose sight of the relationship between the sudden reversal of the deflection and the general physiological state of *O* at the time. From a totally calm, inactive state during which less current was passing through *O*'s skin contact, we have a sudden transition to the opposite pole, that of tension, alertness, strife for activity inhibited with more current passing through the

circuit as the galvanic result. Had *O* been in a mental and physical state of more readiness—for example, less relaxed and looking up at the window when the curtain flopped—it is probable that the rise at *t* would have been more gradual, yet *O* surprised at the unforeseen incident. That may be the explanation as to why *height* as well as *width* of the curve decreased during repetitions of selections, yet the emotional quality remained the same as during the initial presentations. Consequently, we may not deduce that the high rise is due to gaiety or to surprise, but that the one or the other, or some other, is a qualitative accompaniment in the complex of factors in a particular situation.

Likewise, the musical selections provided complex factors as stimuli effecting the galvanic changes. The verbal reports indicate that *O*s appeared to respond to the musical structures which gave gay and triumphant feelings, in most instances, without the experience of anticipation, or prolonged thought. The mental response was momentary, fleeting. The experience consisted in large part of strong urges to sway, tap, beat, march with the rhythm. The total effect appeared to pass over immediately into greater muscular tonus but the actual motor responses were inhibited because of *E*'s direction to *O* 'not to move.' The corresponding deflections coincide with those representing the so-called sensory emotions, such as response to pin-prick or electric shock. All curves of highest rises correspond with experiences or greater physical alertness accompanied with less of the mental tension, study, and anticipation. While the majority of gay and triumphant reactions were those of greater alertness, showing curves with high rises, there are also reports of gay and triumphant experiences accompanying tension, study, reflection, with corresponding flatter curves. The differentiating factors, as indicated by the verbal reports, involve a number of aspects about the musical structure simultaneously presented in the stimuli. Also, with the melancholy and tragic selections,—while the majority showed the flatter rises, there were spots in those selections, which stimulated greater tension, showing corresponding curves with the higher rises, yet the melancholy and tragic moods prevailed.

The data throughout this experiment indicate that any sudden transition from a more languid bodily state to one of considerable physical alertness, regardless of the emotional quality, is an important factor in effecting a high rise in the curves. It is very probable, as the data tend to indicate, that certain affective qualities will accompany this sudden change in physiological adaptation—and others not. Our common experiences, for example, reveal that gay and triumphant music would stimulate tension and alertness in a manner more similar to that which effects laughter, surprise, and startle than to that which effects tragic, sorrowful melancholy qualities. But the fact that there are types of situations which normally would be expected to elicit certain qualities over others would not justify our concluding that those specific qualities would be the ones actually experienced by *O*. We have to consider, always, the total set of conditions. While the data show that certain affective qualities accompanied suddenly increased tension more frequently than others, there were marked variations determined by *O*'s previous conditioning. It appears, therefore, that when the total situation-response is understood, the height of the curve may indicate, along with other factors, some emotional experience lying within a certain range or group of affective qualities. This is as far, in relating height of the curve to qualitative differentiation of emotions, as our data will permit us to go; it appears correlated with a complex of factors not definitely analyzable by psychogalvanic-introspective technique.

Summary. A. Galvanic Data.—The galvanic skin reflex presented three general characteristics; their possible significance and application to the analysis of musical appreciation, as revealed in this study, are as follows:

1. *Amount* of change in resistance corresponded positively with degree of affective tone experienced. Selections which elicited deflections above *O*'s mean galvanic responses are those whose corresponding verbal reports indicated the stronger affective experiences.

2. *Time* of change in resistance. Selections were differentiated as to whether or not they (*a*) effected sudden transi-

tion from a relaxed to a more tense, alert state in *O*, (*b*) set off anticipation, contemplation, mental strain, analyzing attitudes, (*c*) effected relaxation, with or without mild affective responses. R-changes corresponding with the first type were 'rapid'; those with the last two, 'slow.'

3. *Kind* included decreased and increased R-changes, the former coincident with increased alertness and tension; the latter, with increased relaxation, rest, quietude. According to the protocols, both corresponded with affective and non-affective responses.

No one aspect of musical structure appeared causally related, singularly, with any one of these characteristics; on the other hand, the protocols indicated that a complex of musical factors, not definitely separable subjectively, nor analyzable by psychogalvanic technique, connected in an overlapping manner with each characteristic.

B. Subjective Data.—Predominately, sombre music with soft, rich melody and harmony induced relaxation, with moods of calm, soothing, peaceful qualities; gay music of rapid tempo induced alertness, with moods of joyful, cheerful, frivolous qualities; tragic music induced alertness, with moods of anxiety, conflict, struggle; martial music induced alertness, with moods of triumph, force, power, courage, hope.

Based upon frequency of mention, melody, harmony, rhythm, design, and tone color, in the order given, are the elements of music most enjoyed from the selections used. Design, when intricately interwoven, peculiar, or very inconspicuous, inhibited æsthetic response and induced inattention, for the musically *untrained* students; but usually quickened attention and interest toward the selection for the musically *trained* students. Selections uninteresting in design, harmony, or melody induced mild displeasure and indifference; no selections induced displeasure because of rhythm or tone color. Intense displeasure was not experienced from the music per se in this study. Indifference to each type of music was induced by re-presentation of the selection within 15-minute intervals; affective responses were reëstablished after the two-week intervals.

Conclusions.—Results of the present study indicate that the psychogalvanic reflex was of little value in the specific analysis of musical appreciation; and that, at its present stage of development, galvanometric technique for psychological application and interpretation is unprofitable for the following reasons: (1) The impossibility of isolating and controlling all of the phenomena and conditions involved; (2) The intangibility and variability in the subjective aspects involved in responding to complex stimuli; (3) Lack of sufficient knowledge of the true physiological significance of the reflex; (4) Inadequacy of present mathematical formulas for treating scores expressed as ratio values.

(Manuscript received November 21, 1932)

THE EFFECT OF VARYING TIME INTERVALS BETWEEN ACTS OF LEARNING WITH A NOTE ON PROACTIVE INHIBITION*

BY A. H. MASLOW
University of Wisconsin

In the course of a previous experiment,[1] a short rest between the presentation of successive learning problems appeared to cause an increase in the efficiency of learning, and the longer the rest period was, the greater this gain in efficiency seemed to be. The results of a further experiment on the effect of varying the time intervals between acts of learning are described in the present paper. Memory span material was used for learning, thus making possible a large number of trials in a relatively short time.

Procedure.—(1) The apparatus used has been described in detail in the article referred to above.[2] A screen (5′ × 3′) with a window (6″ × 1″) in its center was placed on a table, the subject and experimenter sitting on opposite sides of the screen. The material to be learned was exposed to the *S* in the window.

(2) The learning material consisted of 100 lists of nine meaningful but unassociated words each. The words were picked by chance from the list of 'Familiar Words for Learning Experiments' which Cason used in his study of backward association.[3] Several precautions were taken in grouping the words. Words beginning with the same letter or rhyming with each other were not placed close together. All pairs or groups of words with obvious and ready interassociation were broken up. We attempted to make all of the 100 nine-word series equal in difficulty.

* I wish to thank Professor Hulsey Cason for several helpful suggestions.

[1] Maslow, A. H., The effect of varying external conditions on learning, retention, and reproduction, J. EXPER. PSYCHOL., 1934.

[2] Ibid.

[3] Cason, H., Specific serial learning: a study of backward association, *J. Exper. Psychol.*, 1926, **9**, p. 202.

(3) Each nine-word series was exposed for 10 seconds, and the *S* had just enough time to read the series twice. In each trial, he was instructed to read the series twice rather than to attempt to learn only a few of the words. At the end of the 10-second learning period, the card was withdrawn from the *S*'s view, and a blank white card appeared in its place. The *S* was not allowed to recite at once, but was required to wait for 15 seconds, and during this 'retention period' he was instructed to look at the blank white card. At the end of the retention period, the *E* said 'ready' and the *S* was given 10 seconds to repeat all the words he could remember. This last period was called the period of reproduction. The score on each trial was the number of individual words correctly reproduced, disregarding the order of the words recalled. An experimental period of 40 minutes was used for each *S*. Between each of the successive acts of learning-retention-reproduction, there was a delay which was varied according to a definite plan. The present study is concerned with the influence of this on subsequent learning, the delay being the period from the end of one reproduction period to the beginning of the next learning period. The periods of delay or rest that were used were five, ten, fifteen, twenty, twenty-five, thirty, thirty-five, and forty seconds.

The series were rotated in such a way as to eliminate the factors of position and similarity. With different *S*'s, each card was used approximately the same number of times after each delay.

Forty-six *S*'s were used, all being students in elementary psychology. A total of 1137 series was learned by these *S*'s, or an average of 25 series for each *S*. The writer was the experimenter for all *S*'s.

Results.—The principal results are given in Table I, which is self-explanatory. The mean number of words recalled after a delay of five seconds was 4.77 ± .09 out of a possible nine, and the mean number recalled after a delay of 10 seconds was 4.80. In the last column are presented the combined means obtained by averaging the means for 5 and 10 seconds, 15 and 20 seconds, 25 and 30 seconds, and 35 and

40 seconds. These give us a clearer and more reliable picture of the trend of the experimental results.

It would appear, upon examination of the combined means, that we have an order of increasing efficiency, as there is a consistent trend towards higher recall scores for series learned after the longer delays. The differences between the uncombined means do not show as consistent a trend, indicating that a difference of only 5 seconds between delays was not great enough. The differences between successive means (5 and 10, 20 and 25, 35 and 40, etc.) are not statistically reliable, probably for this same reason. The combined means show greater reliability, as we might expect. The critical ratios are as follows:

Combined means 1 and 2 ..... 3.11
Combined means 2 and 3 ..... 1.00
Combined means 3 and 4 ..... 0.63
Combined means 1 and 3 ..... 4.00
Combined means 1 and 4 ..... 4.56
Combined means 2 and 4 ..... 1.63

These critical ratios indicate that the method used is capable of demonstrating statistically reliable differences.

TABLE I

THE AVERAGE NUMBER OF WORDS REPRODUCED IN TRIALS PRECEDED BY DELAYS RANGING FROM FIVE TO FORTY SECONDS

Preceding Delay	Mean ± P.E.$_M$	N	Combined Means ± P.E.$_M$
5 seconds	4.77 ± .09	146	4.79 ± .07
10 seconds	4.80 ± .09	145	
15 seconds	5.12 ± .09	147	5.07 ± .06
20 seconds	5.02 ± .08	148	
25 seconds	5.27 ± .09	135	5.15 ± .06
30 seconds	5.02 ± .09	134	
35 seconds	5.04 ± .09	135	5.20 ± .06
40 seconds	5.36 ± .09	147	

This fact is especially interesting in view of the fact that memory span learning is relatively simple and stable. As I have pointed out in the paper previously referred to, in which the same material was used, reliable differences are not to be expected in a function which is so stable for the individual and

so variable between individuals. The first factor makes for small differences between the means for various delays, while the second makes for large standard errors of the distribution. The writer had not expected to get critical ratios of the magnitude that have been presented. This consideration argues for an even greater importance to be attributed to differences in time relations than our figures would superficially indicate.

Introspective reports taken after the subject had finished showed that the factor of set for the task was important. Many subjects reported that after a long delay they weren't as well set for the task, and they preferred the shorter delay. Our results show that this factor of set was outweighed by other tendencies working in the opposite direction.

The results of this experiment throw an interesting light on the problem of retroactive inhibition. Experiments on this problem usually measure the deleterious effects of interpolated material on previously learned material. The experiment here reported would seem to indicate the possibility that the originally learned material in a retroactive inhibition experiment would also have a deleterious effect on interpolated learning. In other words, the process may be not merely retroactive inhibition, but rather mutual interference. This possibility should certainly be canvassed.

(Manuscript received December 30, 1932)

THE RELATIVE DIFFICULTY OF NONSENSE SYLLABLES

BY WM. C. F. KRUEGER

Colleges of the City of Detroit

This experiment was performed to determine the relative difficulty of nonsense syllables, based on the responses from a large group of subjects. A total of 586 college students, composed of smaller groups ranging from 10 to 47 persons, participated. No subject responded to all of the 2183 syllables used. The relative difficulty of each syllable was calculated from two hundred responses made to that particular syllable. The subjects practiced on one hundred syllables to familiarize them with the procedure and the problem involved. These instructions were read and repeated or explained in greater detail when requested.

"This experiment is to see what meanings or ideas are associated with nonsense syllables, or are aroused by them. On each paper before you, you will find a series of numbers, 1 to 50. A short and a long line follows each number.

"I shall spell each syllable twice; for example, B-A-K, . . . B-A-K." (The letters were pronounced so that the total time for spelling, including a brief pause between the two spellings, was about four seconds.)

"After you have heard the syllable spelled for the first time, you PRINT each syllable, as spelled, in capital letters on the short line." (Examples were demonstrated on the blackboard.)

"Within the few seconds allowed, before I start to spell the next syllable, you are to make a notation of the idea aroused by each respective syllable on the long line. This notation may be a word or a phrase.

"If no meaning comes to you at once, that is, before I start to spell the next syllable, make a dash or draw a line in the place of the notation, or write the word 'none.' If the syllable should suggest the word 'none,' please put that notation 'none' in quotation marks so that I can distinguish between no meaning aroused by the syllable, and the meaning of 'none' aroused.

"Obviously you cannot get a meaning for every syllable. You are not expected to find a meaning for each syllable. It is impossible. However, do your best to follow these instructions.

"You are not to go back and fill in a meaning after the period allotted to that syllable has passed. Neither are you to change any notations. If you should happen to think of two or more ideas for any one syllable, put down the first idea that flashes into your mind. Do not erase any answers. Your first idea is to stand as recorded."

A total of seven seconds was allotted for each syllable; that is, within this period the syllable was spelled twice by the author, the subject printed the syllable and wrote down the requested notation. During each sitting 300 syllables were presented. The subject participated in four sittings. No additional practice periods after the first meeting were given. The syllables were presented in a counterbalanced order according to definite plan which eliminated the possibility of undue practice effect for any syllable.

The following list groups the syllables according to the frequency of meaningful responses.

TABULATION OF FREQUENCIES OF RESPONSES

Relative Frequency	Syllables Included
1.000	DOC DOL GAP GAB GUS HAL KAT LAB LAF LOC LOV LUC MAC PUF PUL RAH REM SOL SUP TAM TIL ZIP
0.995	DEB DOM FEL GAR GIV HIL HOM KEN LIT PIL RAZ RUF SAC SOM SOX TEL
0.990	DUM FAM HAP HEL HUR HYP JAN JIN KIL KIS LAK LAM LAS LUX MAS MID MIN MIS MUS NAT NIT PAS POL SAX SED SEP SOK TAK VET YAS YEL
0.995	BOM CUS DEM DUK DUL GAD KOP LIM MAX MED MON NOS PAK PIK RAK REP RIT ROL SAK SAV SUC SUR TOL
0.980	COF DOR DOV DUS GAL GEN HAW HOL HON HUL HYD JAC JEL JIL LEW LIF LIV LOF LON MAL NEC NOK RUD SAN SOC TAC TIC TES TUC TUN WYN
0.975	BAL BEL BUL CAS CUF DET DIF DIS DIZ FEM FUL GYP HIK HUS JUS LAQ LAZ LOK LOS MAK MES MOR MUF NIL PEK SEN TAF TEX TIK TUM VIC VOL
0.970	BAS BON CIS COL COM CON CYN DEL DIX DOP FAL FUS GEL GIT GOB GON GYM HAN HAV HUP JAG JAZ JUN KAY KOT LAC LAN LES LIC LOB LYK MAZ NIC NIG NIK PAC POX ROC RYP SAL SID SOR SYN TOG TYP VAC WAL
0.965	CIG CIN COR COZ DOK FUZ HEC HIC JAS JIF JON JYM KAP LIK LUR MAG MEX MIL MOD PEC PUG RAF REV ROP RUS SIK SOF SYD SYL TYM VAL VAS
0.960	BES BOL BUZ CAV CUL DAF DAZ DUC FOS FYR GAM GIM JAK JOC JOK KAN KOW LEP LUV NEK NOC PIC PUR QAK QIT RAV SER SUD SYR TAD TUX WIC WIL WUN

TABULATION OF FREQUENCIES OF RESPONSES—*Continued*

Relative Frequency	Syllables Included
0.955	BIZ CEL DAR DEP DIK DIV DYM GIP GOP GUL HOR KAR KOS LEF LIN LIQ LUK LYP MEL MOC MOL MYT NES NEV NOC POK POR POW RAC RET REY RIV ROM ROS ROX SIM TER TOK WAT WEL WIP YEZ
0.950	BAW BOS CAF COV FEN FOL GOV HAZ JES JYN LIS MIC MIK MOS MOT NEL PUC PUD RAD RAS REL SEZ SIL SIZ SUF SUK SUL TOS TUK VIK
0.945	BAH BAK CIT COS DAT DEC DIL DOZ FID GAZ GIL JEM KAD KUS LYR MUN MYS PON REG RIF ROK SIG SYM TOR WIZ WOT
0.940	BEK BOR BYD CUM DOW DYN DYP FAK FET FIL FUC HAB JEF JYP LEV LIZ LOD LUF LYN MIR NOZ POS PYL RES ROG ROV TYL TYN VAG VER WAV WEN WIS
0.935	BAC BAM BIL BUC COK GOL HAK HOC HUD HUF JER JOL JOV JUL LAT LOR LUS LYF MOV PYN SEC SEV TID TOX TUF VEN VIX VOT WIF WOR YEP
0.930	BOD BUX CIL DAL FAS FER FIZ GOR HAR HOD JEN KAL KET KUR KYD KYL LUN MAH NER NUF NYT PES RAB REF RYM SIC TAB TAL VIN VUL YEW
0.925	BUR DEV DYC DYK FON FYN GAT GER HED HET HUZ HYM JUR KED LAV LEN LIR LYD LYS PED PIF POC SEY TYR VIZ WER WOM
0.920	BEC BER CYT DUR FAB FAC FAV FIV FYB HIX KAM KUT LOQ MUR MYR PEY PIS POM PYG RYD SUT TYK VEL WOP
0.915	CEN CUZ DAV DEF FAX FIC HIR HOK HOX JAV JYL KIP KYT LEY LIB LYT MOQ NAV QIK RYT SUB WAP WIM ZEP
0.910	BIR BOT BUF DER DIC JIP KIC KOL LYX MER MEY MUC MUK NAS NAW NEY NOB NUG PIR RAQ REB REC ROQ SYK TAQ TYC VIS WES WUZ YON YOR
0.905	BAX BIK DAG FUM FYT LEX NAC NEG NUM PAM PYC QIS RAJ REK SAB SAR SIF SYT VES WAF WAM WIK ZYP
0.900	BEV DEK DES DUF GIB HEK HOS HOV JUM KUP LIX MEG NOM NYL PYT QAC RAX SEM SYP WID WOL WYZ YEN
0.895	BAF BOK BUK DAC DAP DAW DEX DOS DUX GES HAC HAF KIM KOM KOY LYC MAJ MYN NUL NYC PEB PIQ QIL SIQ SOV SUG TAY TIR TIX VAP YUM

TABULATION OF FREQUENCIES OF RESPONSES—*Continued*

Relative Frequency	Syllables Included
0.890	CAK CYS DIR FOD FUG GUD GYL HAX KER KOF KOR KUM LIG LUT LYM NAK NUW PUM QAR ROZ SAQ TEM TIF TOF VIR VYM WAC WEY WOD YAP YIP
0.885	CYD DEY DYV HUX JOQ KEP KYS MIZ NOP NOV PAQ RIN TIS VIT WAK WYF XAM
0.880	DYL DYZ FEV GOS GUZ HEV HIQ HOY HYS HYT KEL LAR LUB MEK NIF PIX PYX QYK RIC SEG SEL TEK TOC VIL VOY WYL YAH
0.875	BYR CIV DOQ FIS FUD GAV GOM HEP HES HIN HOB HOQ JAX KAB KAF LOY LYV LYZ MOK MUZ MYL NYP PYR SOY TAS TUR WIR WIX
0.870	CID CYL FUK FUX FYL GIR HEF JAQ KUF LEM LIY LOX LUQ NEX RYN SEK SIV SOH SUH SUQ TIQ TOB WYT
0.865	BIM CER CIR DAS DUT DUV DYR FES FOT FUT GIF GUV HAQ KAW KUZ LER MUL NYX NEP PEL REN REQ SAF SOQ TEC VOM WEG WOK WUF YUL
0.860	BYN HEX HOF HYK JOS JUD KON LYB MYX NEB NUR PAV RUH RYG SEQ SOG SUX TEP TIY WEX YEG YUR
0.855	BYC BYL CAG CAY DIT DUP DUQ FOW GEY JAD JOW KYN LUD MEC NAH NAM PAG QAY QIR RIS RIZ RON YEH YOK YUS
0.850	DOX FUJ JAL KOX MAQ NEM PAH PUB QES RYS RYZ VOG VAR WYD YAL YOW
0.845	BYF CIK GYR JIB JIT LEZ LUP MEZ NAR ROF RUP SAZ VAK VAZ VIG
0.840	FYG HIZ HYL KUL LEK LUM MOG MOH MOZ TUS TYX VEG WOV WUM YER YID
0.835	BYG CED FOK HIV HOZ KAS NEH NOG PAX PUK PUQ PYK RIK RIL RUL SYF TYD VYN WYG ZAT ZOW
0.830	BAQ BIC FYX GAJ GIS GQF GYN JUY KES KOB MOF NOL QIC VAX WAH WOH YAR
0.825	BUJ DOB DYG KUD QER QOT QYT RYB SYV TUL WUR ZIG ZUP
0.820	BEF BOC BYT CIM COH DYX FAZ JEZ KIQ KOZ MYK QIS ROH RYL SYB TEQ TIG VYC
0.815	CEZ CYR FEB FUQ HEG HUK HYR JUW LYQ NEQ NOX SAH TEY VEW WIV
0.810	BIS CEM CET FEX FYZ JIX KEW KYP LOM MAB POZ QUT VEM VOX WYP ZAP

TABULATION OF FREQUENCIES OF RESPONSES—*Continued*

Relative Frequency	Syllables Included
0.805	DIQ DUZ GAN GOW HUC JOR KAV KEM KIX KYZ MYD RAL RYF RYV SYX WEJ WUD WUS
0.800	BIF HEB JOH JOP KOD KUN LOH MUV NAZ NIB NUJ NYK SYC TEW WEP
0.795	BYZ CEW CUV HIY HUV KUB NIR NUX PUZ PYF QYC RIQ TYF XYZ YAT YIS ZAW
0.790	DIY GUF GYB JEK KIB LEC LOZ NIQ NOY POY QAL QYL SIY VON WAQ WYC YAZ ZAR ZIT
0.785	CES CEY GUR MUJ NYB PYS QEL SIH VOC
0.780	BOH BYM CUK CUN DYF DYT FYD FYS GAF HEQ KIR KIZ KYX LYG MIY PEQ QAX QID RUJ TIB TUY WEV WYM WYR YAR
0.775	BEZ BIY CIQ COQ DEQ DUW FOY GOY HYG JEP JYG JYT LAH LUZ MUQ MUW NUC NUS QUZ. REZ VIP WAB WUL ZAC ZUN
0.770	BOZ DIB DOF DOY FAQ FOV GID HIB HYQ JUB LEH MIG MYZ NYR QAN RYH TYH WAZ ZAK
0.765	CYK DYS FAW FIQ HEJ MOX QIP RUK TAW TUQ VYX WIQ WOF WYS ZIS
0.760	BIQ CEP CYP DYP FUH JUK JYF KUV NOQ QAT TUP VAD VAQ VOS WEH WIY
0.755	BOV CEB FIP FUW GEV GYT GYC HYC JAH LEJ MUP PIV PYQ QEZ QOD TUH VAY VIW WUY WYK WYQ
0.750	CYM DOH DUY FOM HYZ JUZ PEX PIM POH QIN RIQ SOJ SYZ TUW YUH
0.745	FOZ GYV HIJ JIY JUV LOJ MAV MIQ NIM PAZ QYR REW TEG ZEL
0.740	CAH CIP DAH DUJ FIJ FYV GOH JIS JUP JYR KOG NIV NUK NYH PYM QIX RYJ SIB TOH WOS WOZ XOT ZON ZOR ZUM
0.735	BAJ BYK BYS CEV DIH DOJ GEW HUQ JIW KYR LUW MUX NIH NOH NUD PUJ QUL ROJ SEH VOK XEL YOS YUZ ZEB ZYN
0.730	CAX FOC HIG JAF JOD KOC MOY NAL NUB NUP PIJ PIW POJ POV QYD QYN REH SUW WUK
0.725	DIW GAC GOZ HUJ JED JOF KUX MEB NEF NYG POQ PUH QED RIY WEK WUP ZAG ZED
0.720	BAP CAJ DAK GOK JOZ LUH LUY MYC MYQ PUW QET QOK RUW SAJ SEJ TYG VUN WEZ XAV ZAD ZOL
0.715	BAZ CUY DEG DEH JUH KAG KOV LIH NAX NYS QIV QOZ QOY VAM VYK WOB XIP YIT ZAM ZES

TABULATION OF FREQUENCIES OF RESPONSES—*Continued*

Relative Frequency	Syllables Included
0.710	BUQ FUY GEF GEZ HEZ HYB NEZ SIW SUV TAV TYZ ZIL
0.705	BOQ CYZ FEY FIH LAJ MUH NAD NOF NYM QAF QEY REJ VEY VIQ WUT ZAY ZOK
0.700	BOJ FAH FEC GUP HUY KAH KYM MIH MUY RIJ RYK TAJ TEH ZOP ZYS
0.695	BOP CEG CEX CIF JYS LIW MIJ QAD SUY TYS VOP VOZ ZAQ ZIX ZYL
0.690	FEG GUB HUW JOX JUC JYX QAM QUS SOZ TIH TOZ VIH WEM ZIN
0.685	BAV CUH DYQ JEC NUV NUZ PEM PUX QUC QYZ TEB TIZ TOQ VOD VUZ VYL ZAV
0.680	BEQ DYW GYF GYS JIZ KOH MYH NUH QAV SYG TUD TUV TUZ ZOX
0.675	BUW CIX DEZ FIY GAW GIW GIZ QOR QUR VEC WYH YAC ZEW ZOM
0.670	BYH GYD GYW KAC KAZ LYH NYF PEZ QEW TYQ WEF WOG YUN
0.665	BUH CIZ JEG KYC PEV POG SYH TYB VUM WYV XEM YIR
0.660	CEK FEJ HYX MYG PIY PYW QEK QOF QYP QYS SEF YOT ZEY ZUR
0.655	JEY JIR JOM KUC LYW SUZ TEZ VYP VYS WOY XIT ZAN
0.650	BEM FIK GAK JYB KIG NYQ PAJ PID PUY QEP RUC RUX TIW VEP VYG VYT WOX XEN XES ZAL ZEF ZOG ZYT
0.645	BEP BEW DIJ FEK LEQ NAQ QAD QIB VAH VUG VYF ZID ZIM ZIR ZOD ZYM
0.640	BYQ CAZ CEF CIB CIY CYV FIW FOQ GEB GED HAJ HYV JID JUQ KEQ KUH PYZ QAW YAN YEM ZOT
0.635	CYX JUX KIH KUQ MEH MEJ MEQ NUQ NUY NYW RIX SYQ VIB YOD YOL YOM YUW ZUK
0.630	BEH CUQ GEP GIY NIZ NOJ NYD PIH PYV RYC VEZ WYX YOH ZEX
0.625	FIM FOJ FYK GAH GOC HYN KEB LEB NYV TUJ VOR YUP
0.620	BYV FYM KAX KEH MUB NAF QUD TAH TOV WYJ YAX YIN YOB YOC ZER

TABULATION OF FREQUENCIES OF RESPONSES—*Continued*

Relative Frequency	Syllables Included
0.615	CUW CYG GAQ HIF HYJ JEH NIW RUZ VEK VUR WIJ ZUC
0.610	BUP CUG JAT MEP PUV SUJ VEH VIF WEC XAG XET YIZ ZEM ZIK
0.605	DUH MIF NID NIS PAB QAB QUF WOQ ZAX ZEK ZIF ZIV ZOH
0.600	BIV CYB CYF FAP JIQ MOJ QOM QYS TAZ VIY VUT VUY VYR XUM YEX ZUB
0.595	BYX CAQ FYJ KAJ LUJ NIY PEH PIZ PYJ VYD WIH WOC XAY XOR YUG ZOC ZUL
0.590	CUX JEB KEC KIV MYW QYV YED YUK
0.585	BIJ GOJ HOJ JEX QUP RYQ RYW VUC WIB XOS
0.560	BIX KUY MIB MIP NYJ POF QIM VEB VOJ VOQ VUS WUX WYB YUT ZEN ZIC ZUD ZUT
0.575	BEJ CUJ GUH KAQ KEF KEX QAP RIH RUQ RYX VAF VEJ VEQ WAJ WUH YIL ZYG
0.570	BIP CIW KUG KUW KYB MAF MYF QOH VID XAS XAW ZIB
0.565	JEV MIW MYJ NAJ QAS RUY ZEV ZYC ZYX
0.560	FOH FYK KIY KYG QOL SYW XEP YIF YOP
0.555	GEH KOQ NEJ NIJ QEX XAG XYL YOG ZIH ZOF ZUS ZYR
0.550	BUV FAJ FEH GAX GUJ KEJ POB PYH QEN QIG QUN QYX XER XIK XON XUT ZAS ZUG
0.545	BYJ DAQ GOQ QON VUX VYZ WUB YEJ ZET ZOV ZUF
0.540	GIJ JIV JYK QEG QUK XAN XIL
0.535	FYP QAG TYW VYB XED XOQ YAG
0.530	GUW LIJ PYD QAZ SIJ TOJ VAJ WEQ WUV XAL XIG XYT YUC ZAB ZYD
0.525	BEX GYZ TEJ
0.520	BOF BYW GYX KYW QEM QUV QUY VUW TEF TIJ YAQ
0.515	DYJ FYQ FYW HIW HYW JIC JYW KEZ KIJ MIV SEB TIV WUG YEC YOZ ZEC
0.510	CEQ COJ KYH RIW VAW XUD XUL XUN ZAJ
0.505	FYH GUK QOG RUV XID YAB
0.500	BIW CYW FUB KOJ QYB QOS VED XIN XOD XUS YIM
0.495	CYQ JYZ LYJ QOW XAK XAT YAV YOX ZYF

TABULATION OF FREQUENCIES OF RESPONSES—*Continued*

Relative Frequency	Syllables Included
0.490	CYH FUV GYC KYQ PEJ VUJ XAR YUF
0.485	DAX KIW KUJ NYZ ZAF
0.480	JEQ QEB QEC QUB QUW VOB XEW
0.475	GIX MYP QUM XOY YEF
0.470	BIH KIF VUD VYW XOL
0.465	FEQ KEV SYJ YUB ZEH ZIJ ZYB
0.460	GIK GUC JYV KYF QOC VAB VYQ XOW
0.455	CEH FEP JIK MYB XAP XOP
0.450	JYH VEF WOJ XAB XOC XUR YAF ZYK
0.445	CIJ DAJ JIH QIW TYJ WUQ YUX
0.440	GIC GIQ MEF QEV QIF VUK VUP XIR ZOQ ZOY
0.435	BYP JUF MYV QOB QOP QOX TEV TYV VIJ VUB XOK XUB ZOS ZUH
0.430	GUX PIB QUX QYF YAD ZOB ZYV
0.425	GEC GYK JYC KYV QUH WUC XAD XAJ XIC XIV XUG YIC ZIY ZYQ
0.420	CYJ QYM XEQ XYS ZEJ
0.415	GUQ QIY XOM ZUX
0.410	DYH GEQ GIH JYQ PAF XIF XIM XYP YEK YOQ ZOJ
0.405	GOX QIH XYN YIN
0.400	GEJ GEK PEF QIJ XEC XEK XOH
0.395	QEH QYW VOH WUJ XIQ XOG YUD ZAH ZUQ
0.390	MEV VUF XAZ YIB YOV
0.385	CEJ CIH GEX KYJ QUG VUH XIS XOB YIW ZEG
0.380	FUP JYD QYH YIG
0.375	YUV
0.370	GYJ QAJ XUP XYD YIQ YUJ
0.365	XEZ XYB YEV YIG
0.360	DEJ XIB XOV ZUY
0.355	XUK
0.350	XEY YEQ ZEQ ZUG
0.345	XEG XUZ XYK YIX
0.340	GYQ ZIQ
0.335	PYB QAH QEF

TABULATION OF FREQUENCIES OF RESPONSES—*Continued*

Relative Frequency	Syllables Included
0.330	GYH ZYH
0.325	YIK XUW XYR YAF
0.320	QYG XYC
0.315	XAF XAQ XUC YIH ZIW ZYW
0.305	QOV VOF XOZ XYM ZUJ
0.300	XIH XIZ XYV XYW YEB
0.295	VYH XIJ XOF
0.290	QOJ XUY YUQ
0.280	XEV XYG
0.275	ZUW
0.270	XAH
0.265	XUV
0.255	VUQ
0.245	QEJ VYJ XUH
0.240	YIJ
0.235	XEF XYQ YOF YOJ
0.230	XEH XUF YIV
0.225	QUJ XEB XUJ
0.220	XYF
0.215	XYJ
0.195	XUQ ZYJ
0.190	XYH XIY
0.185	QYJ
0.180	XEJ
0.165	XIW
0.145	XOJ

(Manuscript received December 9, 1933)

THE SENSORY FUNCTION OF THE NON-PREFERRED HAND

BY JOSEPHINE M. SMITH

University of Missouri

I. Problem and Method

This is the second portion of an experiment conducted to test the hypothesis of the sensory function of the non-preferred hand. The suggestion offered is that the non-preferred motor hand is preferred for sensory functions; that is, if an individual uses his right hand preferentially in the motor activities of writing, throwing, cutting, etc., he will probably use his left hand preferentially for manipulating objects to discover tactually their shape or spatial location. *Sensory function* is defined here as any use of the hand in place of a sense organ, for the gaining of information rather than in execution of a task.

As a test of this hypothesis two experiments were conducted. I. Three groups of normal adults, right-handed, left-handed, and ambidexterous, when blindfolded, replaced the shapes in the Goddard form board. The purpose was to discover whether or not the subjects consistently preferred one hand for cutaneously exploring the blocks and depressions. The experiment and results have been reported.[1] II. The present portion consisted in having two opposite-handed subjects read Braille script with their fingers, a sensory task. The excessive time necessary to the completion of the experiment necessarily limited the number of subjects who would be willing to coöperate. The fewness of numbers is partly compensated for, however, by the strict criteria used in selecting the subjects and by the intensiveness of the method.[2] Subject A was right-handed in all activities, uni-

[1] J. Exper. Psychol., 1933, 16, 271–283.

[2] A third subject, who was ambidexterous, also performed the Braille reading. However, her data had to be thrown out because of the huge variations caused by her

and bi-manual; he showed right eye dominance. Subject B was left-handed in all activities, uni- and bi-manual; he also showed right eye dominance.[3]

The two subjects whose data are considered repeated exactly the method as described previously.[4] That is, each subject read a certain amount of Braille six days a week, using alternately the left hand alone, the right hand alone, and both hands together. In every case two days were given to each method of reading, the first one for the direction right to left, the second for the direction left to right. This control of direction precluded the possibility of muscle preference being interpreted as hand preference. Time required for each reading was recorded with a stop-clock. Of course the material for each hand was comparable,[5] and both subjects read the same material. The subjects carried on the reading for thirteen weeks, or 78 practice periods, a length of time which extended long after the learning efficiency had ceased to rise.

II. Results and Treatment of Data

The experimental data are included in a table for each subject showing the time required in reading the blind script in each consecutive reading trial. The data were also recorded on a large graph, reading times being plotted against the number of preceding practices. The learning curve thus presented was negatively accelerated. Then the experimental curve was smoothed by the use of Meyer's arc-cotangent function formula,[6] and the curve thus computed drawn in on the same graph.

When the computed curve was superimposed upon the

irregular schedule and time of reading. During the first few weeks she read every day and at the same time of day; but for the greater portion of the time she read at times of day varying from 7 A.M. to 12 P.M., and frequently let a day pass without doing the reading. These irregularities made her reading data incomparable with those of the other two subjects.

[3] Method of determining handedness and eyedness is given in the report of Part I, to be published in the Jr. of Exp. Psy.

[4] Which hand is the eye of the blind?, *Genet. Psy. Mon.*, V, 1929, 3, 225–227.

[5] Ibid., p. 225.

[6] Meyer, M. F., and Eppright, F. O., The equation of the learning function, *Am. Jr. of Psy.*, 1923, XXXIV, 203–222.

experimental curve it was seen that not only did the experimental points vary above and below the smooth line; they also varied regularly. The points representing reading by the left hand were marked *l;* those by the right hand *r;* and those by both hands *b*. It was then clear that for the right-handed subject the *r* points were most of them above the smooth line; the *l* points were below the line; and the *b* points vacillated a little in either direction. The reverse was true of the left-handed subject.

The ratios of the experimental to the computed time for reading at each sitting were figured for each subject, and then arranged according to the handedness of the subject and the hand used in reading. A summary of these tables is presented below, Table 1. The ratios are to be read thus: a ratio of 1

TABLE 1

SUMMARY TABLE

Average Ratios of Experimental to Computed Reading Times with Probable Errors

Subject	Material	Right Hand	Left Hand	Both Hands
A (R)	1	1.04 ± .09	.95 ± .03	.97 ± .04
	2	1.03 ± .06	.96 ± .02	.96 ± .03
B (L)	1	.85 ± .10	1.16 ± .15	.99 ± .14
	2	.91 ± .12	1.15 ± .13	.95 ± .12

means that the time actually required for reading was the amount expected according to the computed curve; a ratio above 1 means that the time actually required for reading was longer than that expected; and a ratio below 1 means that the time required was less than that expected. What was expected was, of course, the average time as computed by the arc-cotangent formula. In the computation the ratios for each hand in the two reading directions were thrown together, as there was but a slight difference between the directions. The average of the material and direction differences for subject A was .01, while the average of the hand differences was .08; the average of the material and direction differences for subject B was .075, and of the hand differences .26. This shows that neither material nor direction was a factor in the reading efficiency.

It will be noted that for both sets of material the ratio for the right-handed subject's left hand is less than that for his right hand; and similarly the ratio for the left-handed subject's right hand is less than that for his left hand. The more efficient reading hand for A was the left, and the less efficient the right; for B the more efficient was the right hand, and the less efficient the left. These differences are statistically reliable, as shown in the following table. The differences between the efficiencies of the preferred and non-preferred hands are too great to be accounted for by chance; for these two subjects there is a real difference in the sensory performance of the right and left hands, in both cases in favor of the non-preferred motor hand.

TABLE 2

RELIABILITY OF DIFFERENCES IN EFFICIENCY OF THE TWO HANDS FOR BOTH SUBJECTS

Subject	Material	Right Hand	Left Hand	Diff.	S.E. Diff.	$\frac{\text{Diff.}}{\text{S.E. Diff.}}$
A (R)	1	1.04	.95	.09	.027	3.33
	2	1.03	.96	.07	.022	3.18
B (L)	1	.85	1.16	.31	.052	5.96
	2	.91	1.15	.24	.056	4.28

The ratios of the reading with both hands fall in every case below 1.00. That is, both subjects read better with both hands than would be expected, and also better than with the preferred motor hand; in two of the four cases the difference between the preferred hand and both hands is reliable. The ratios for both hands are greater than those for the non-preferred motor hand in three of the four cases, and equal in the fourth. None of these differences is reliable. The fact that both hands read better than the preferred hand alone, yet worse than the non-preferred hand alone, is very probably due to the fact that when using two hands the subjects tended to do most of the work with the hand for which the reading was easier, and to drag along the other. Comments made during the experiment by the subjects support this idea. One of the subjects (right-handed) remarked that when reading with both hands his right hand slowed him up, and that he

could do better if he were allowed to read with the left hand alone, using the right hand simply to hold the paper.

III. Discussion

The small number of subjects used in the experiment prevents the drawing of generalizations. However, the data are not only statistically reliable, but are also consistent with findings in previous experiments conducted on the same problem. In the preceding study which prompted this one the subject, right-handed, read more quickly with the left hand and more slowly with the right. The right-handed subject in the present experiment also read more quickly with the left hand, and the left-handed subject more quickly with the right hand. This difference is supported by the results of the first part of the present experiment, that the left hands of the right-handed subjects, and the right hands of the left-handed subjects, were more frequently used in the sensory handling of blocks and depressions. Surely this use of the non-preferred motor hand in the sensory capacity is not the result of chance in the two types of activities studied. It seems probable that there is an underlying causal factor related to the general phenomenon of motor and sensory sidedness which is responsible for the type of activity observed. The nature of this factor cannot be guessed at this point, but from the results of the block performance it appears unlikely that the sensory use is the *result* of the motor use. It is more likely a concomitant phenomenon, both types of hand preference resulting from a common cause. This has been more fully discussed in the previous article.

IV. Summary

This is the second portion of an experiment conducted to test an hypothesis formed on the basis of a previously conducted experiment. In the original performance a right-handed subject read manually some blind-script material. The shorter reading time of the left hand was interpreted as being due to a preferential use of the non-preferred motor hand in sensory functions. Part I of the second experiment

consisted in having right-handed, left-handed, and ambidexterous subjects when blindfolded replace the blocks in the Goddard form board; the results showed preference in manual exploration for the non-preferred motor hand. In the present portion of the experiment a left-handed and a right-handed subject each read the blind-script used in the first instance. The results are consistent with those previously obtained and statistically significant. The right-handed subject showed a shorter reading time with the left hand, while the left-handed subject showed a shorter reading time with the right hand.

This consistency in the more frequent (as in Part I) and more efficient (as in Part II) use of the non-preferred motor hand in sensory activity is interpreted as being due to a preferential use of that hand for sensory functions.

(Manuscript received October 18, 1932)

SELECTING AND TIMING APPARATUS*

BY F. H. LUMLEY

Ohio State University

In certain serial-learning problems it is desirable to present a number of choices to the subject at each point in the series and have the subject select the right response by trial and error before proceeding to the next step or point. The ordinary maze, with its choice between the true path and cul-de-sac, is this type of a problem. Apparatus for the electrical control of such "multiple-choice" series has been described by Henderson,[1] Lumley,[2] and Esper.[3] Indeed, the mechanical testing device designed by Pressey [4] for use in the classroom on examinations with multiple-choice questions can be used for this type of learning experiment.

The essential part of such apparatus is a "stepping selector" which will change the set-up each time the subject makes the right response. Suppose the subject can make eight possible responses. The electrical connections are adjusted so that only one of these responses will make electrical contact. This electrical contact when once made will act on a magnet and ratchet device so that another response will become the right one and will in turn make electrical contact. As ordinarily constructed, these stepping selectors do not move forward to the next setting until after the previous right response has been completely made, *i.e.*, the electrical contact broken. Stepping selectors are sometimes made in rotary style after

* Help in devising this apparatus was furnished by Dr. Renshaw, Mr. Hampton, and Mr. Okey.

[1] Henderson, A. L. "The Design and Construction of an Electric Maze." M.A. Thesis. Ohio State University, 1923.

[2] Lumley, F. H. "An Investigation of the Responses Made in Learning a Multiple Choice Maze." *Psychological Monographs*, 1931, XLII, No. 189, 61 pp.

[3] Esper, Erwin A. "Polytactic Manual Movement Apparatus." JOURNAL OF EXPERIMENTAL PSYCHOLOGY, XVI (February, 1933), pp. 161–174.

[4] Pressey, S. L. "A Simple Apparatus Which Gives Tests and Scores, and Teaches." *School and Society*, XXIII (March 20, 1926), pp. 373–376.

the manner of the Ranchburg tachistoscope, with electrical contacts set against the face of the disk. Telephone companies have used a stepping device to change four contacts at each step.

A stepping selecting device would possess great advantage if the pattern of right responses could be altered very quickly. This would permit the apparatus to be changed rapidly from one experiment to another or allow the experimenter to adjust the learning problem to each subject. There are also occasions where no repeated pattern is desired, but rather an endless variation in the sequence of right responses. This variation can be obtained if the pattern is repeated at long intervals. If able to function in this way, the apparatus would find use in "pursuit" or association experiments as well as in serial learning problems. Such requirements call for an electrical stencil. In other words the pattern must be punched into a card which will make the proper electrical

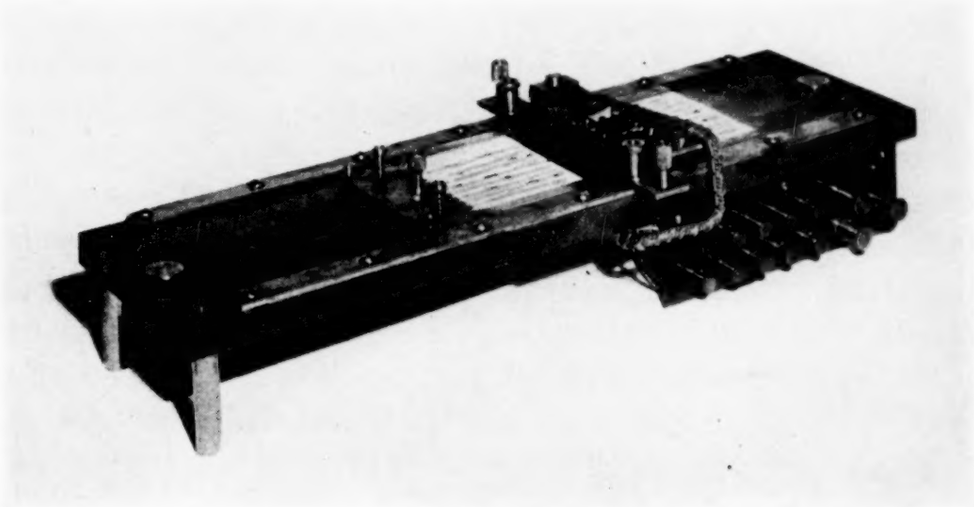


PLATE I
Stepping selector apparatus

contacts, and which may be replaced easily by another different card. To use the pattern over again, the same card may be started at the beginning. To obtain an almost endlessly varying pattern, only a large number of different punched cards is needed.

In Plate I is shown an apparatus designed to use the

Hollerith punched cards as stencils for electrical contacts. The advantage of using these cards comes from the fact that they are especially planned for use in electrical contact tabulating machines, and they may be quickly and accurately punched on standard Hollerith equipment. This last advantage is a most important one since the use of stencil devices has suffered because of the great labor required to accurately cut each stencil. It does not seem worth while to give the exact specifications for the construction of the apparatus shown in Plate I because there are certain improvements needed in its design. Furthermore the Hollerith equipment includes a duplicating device which might be adapted to the same purpose as this apparatus. For these reasons only the general features will be described.

This punched card stepper selector consists of a heavy base block guiding and supporting a moveable carriage. The carriage rides dry upon the block and the bearing and guiding surfaces must be very carefully milled to allow the carriage to move freely back and forth and yet be held accurately in place. At each end of the block are eccentric stop pieces to limit the movement of the carriage. The carriage accomodates one card of the Hollerith type (7 3/8 inches by 3 1/4 inches). This card is slipped into guides on either side of the carriage and held firmly by a spring clamp at one end. The guides and clamp insure the accurate placing of the card. As stated before, the holes in the card are punched with Hollerith punching equipment. They are small rectangular holes about 1/8 of an inch high and less than 1/16 of an inch wide. Lengthwise the card provides for 80 columns with one hole in each column. In width the card has a capacity for 12 rows; 10 are used for this apparatus. A bridge of composition insulating material extends over the carriage and on it are mounted the ten spring bronze wire contacts. The fingers of these contacts are held laterally by slots in the composition bridge and are adjusted to slide easily over the punched card coming exactly into contact with the punched holes. The alignment of ten contact points is a delicate matter when contact must be made within 1/16 of an inch. Brush contacts, such as

used in the Hollerith equipment, would undoubtedly be more efficient. To obviate any chance of failure with the spring bronze contacts because of poor alignment, double consecutive holes are punched in the cards used for this present apparatus. This gives a hole somewhat over ⅛ of an inch wide, which is sufficient to ensure contact even when the spring bronze points become slightly disaligned. When the carriage moves in one direction the spring bronze contacts slide over the card coming into contact with the bare metal of the carriage plate wherever holes have been punched in the cards. To return the carriage, the contacts are raised about 1/16 of an inch by moving forward a horizontal plate or bar. This bar can be fastened in either position, permitting the contacts to rest on the card or holding them in the air. Each contact is connected by a lead to a separate binding post and the whole block and carriage is connected to the common lead.

The actual stepping of the device is accomplished by a strong magnet located in a rectangular opening in the middle of the block (not visible in Plate I). Running the length of the carriage in the center is a rack, milled or cut into the bottom or sliding face of the carriage. A catch on the magnet arm moves back one rack tooth every time the magnet is energized and drives the carriage forward by spring action when released. A stop prevents the spring from pulling the carriage more than one notch forward at a time. The number and placing of the teeth in the rack correspond exactly to the number of columns in the Hollerith card and its position when held properly in the carriage. (In the present instance it is necessary to have the magnet step the carriage forward two teeth at a time.) A lever located on the side of the block depresses the catch on the magnet arm and can be used to move the carriage forward one notch at a time. It must also be held down when the carriage is moved back to the starting point. Since it is important to preserve the carriage plate from the effects of sparking when contacts are broken, a relay shuts off the current during the time the carriage is stepped forward to a new position and the contacts are made and broken.

Under present conditions the apparatus permits a pattern with 40 steps to be used. When redesigned, the full 80 steps or columns of the Hollerith card should be available. The pattern can be repeated as often as desired; shorter patterns may be used by punching in a repeated pattern in successive parts of the card. Longer patterns may be used by stopping the device for a few seconds and inserting other cards As the apparatus stands, there are 10 different contact points and therefore ten different keys may be used by the subject in a learning problem. However, one of the contacts may be made to work a relay changing the common lead at the subject's keyboard This will give 18 possibilities instead of 10, and this procedure may be carried still further.

In the experiment for which this stepping selector was used the following requirements were made of the whole apparatus. The apparatus must sound a particular tone for a specified length of time, remain quiet until the subject pressed the proper key, and at the release of the key sound another tone, starting the cycle over again. If the subject pressed the proper key while the tone was still being sounded, the apparatus must immediately stop the tone and be ready to sound the next tone upon release of the key. The stepper selector described above determined which tones were sounded and which keys were the correct ones to press in association with the tones. The 'timer' to be described below determined the length of the tone and permitted the cycle to be interrupted at any time.

Plate II shows the timer used in this experiment. It consists essentially of a large solenoid magnet, a moveable rack, a motor-driven gear, and contact points. In conjunction with appropriate relays its operation in the above experiment is as follows: When the subject releases the right key the solenoid circuit is broken, permitting the lower (horizontal) spring to press the enclosed rack against the revolving gear. This gear revolves at a constant rate and causes the rack to move upward until the insulated top of the rack touches the contacts mounted on the post. During this time the tone is sounded. When the contacts are closed and held through a locking

relay the current is sent into the solenoid and the rack is drawn backward away from the revolving gear. The vertical spring on the rack returns it to its starting position. The locking relay remains locked until the subject has found the correct key, and then after release of the correct key the cycle starts over again. If the subject finds the correct key before

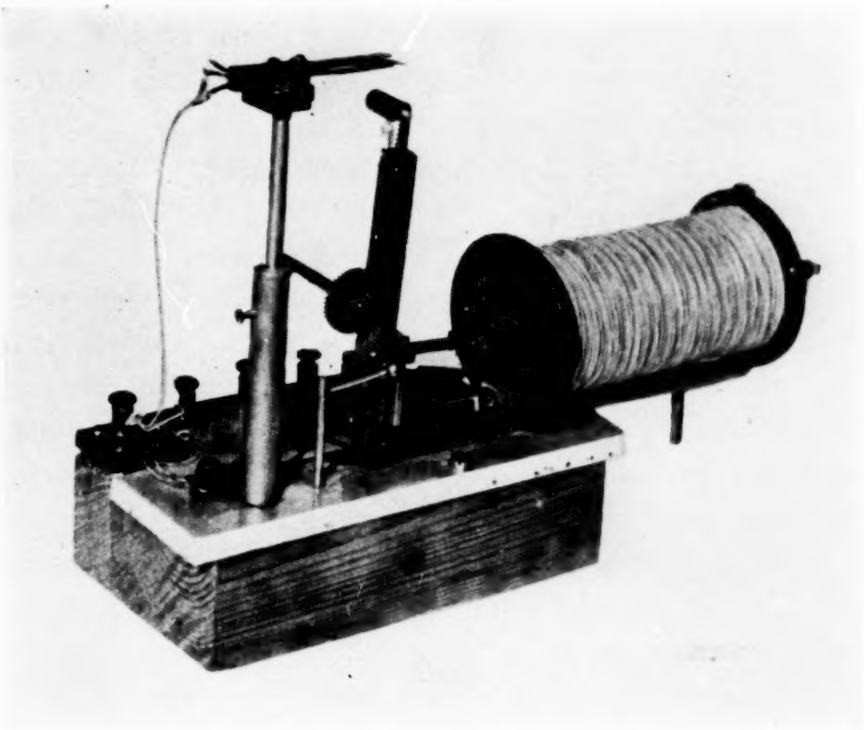


PLATE II
Timing device

the rack has ascended to the upper contacts (before the tone has finished sounding) the solenoid is activated, the rack drawn away from the revolving gear and the apparatus is ready to start over again. It may be thought that this device is not very accurate, since there are many places where variations in timing could arise. Its accuracy depends upon the purpose for which it is used. In this experiment it was certainly accurate enough since variations in the length of the tones as measured by a chronoscope all fell within 1/10 of a second. The length of the tone was 1.6 seconds.

Here again the general ideas of construction will be given rather than the detailed specifications. The moveable ⅛-inch

brass rack is housed in a channel cut in an upright brass bar which is pivoted at the bottom in a piece of supporting channel stock. A brass plate screwed to the upright brass piece holds the rack in its channel. A slit in the brass plate permits the revolving gear to engage the rack when the action of the horizontal spring is unopposed. The moveable iron core of the solenoid is attached by a double pivot to the back of the upright brass housing. Under action of the 110 D.C. current, the solenoid core moves back about a quarter of an inch, which is enough to clear the revolving gear and the upper contacts. At resting position the head of the rack projects but slightly from the housing. On the head of the rack is fastened an insulating piece and from one side of this a long vertical spring extends to a catch on the side of the rack housing. This spring helps to return the rack quickly to the starting place after it is released from the revolving gear. The upper contacts are simple telephone relay contacts mounted on a sliding post. The post is contained within another post and by means of a set screw the height of the upper contacts can be varied. The timing of the device can thus be altered by changing the position of the upper contacts and also the speed of revolution of the gear. This gear is run continuously by a synchronous motor. Two lower contacts are also provided so that the return of the rack closes a circuit. These lower contacts permitted the apparatus to play a whole sequence of tones without stopping in the experiment for which these two pieces of apparatus were devised in 1930.

(Manuscript received December 14, 1933)